CPExpert™

DASD Component

Computer Management Sciences, Inc
6076-D Franconia Road
Alexandria, Virginia 22310-1756
voice: (703) 922-7027
fax: (703) 922-7305
www.cpexpert.com

This manual applies to the DASD Component of **CPExpert**tm, a proprietary software product of Computer Management Sciences, Inc., Alexandria, Virginia, United States of America.

The information in this document is subject to change. Comments or suggestions are welcome, and to the extent practicable, will be incorporated in revisions to this document. Please send all comments, suggested new rules, suggested changes to existing rules, suggestions for improvement to the software or documentation, or any other advice to:

Computer Management Sciences, Inc.
6076-D Franconia Road
Alexandria, Virginia 22310
(703) 922-7027   FAX: (703) 922-7305
www.cpexpert.com

DISCLAIMER

The advice, recommendations, or otherwise contained in this document represent information generally available in the public domain, as contained in vendor manuals, published in articles or papers, presented at professional conferences, or otherwise commonly accepted in the professional community. Neither Computer Management Sciences, Inc. nor its representatives make representations or warranties with respect to the applicability or application of any advice, recommendations, or otherwise, contained in this document or in any results from applying the CPExpert software, to any particular computer system or computer installation.

TRADEMARKS

CPExpert is a trademark of Computer Management Sciences, Inc. IBM, MVS/370, MVS/SP, MVS/XA, MVS/ESA, Enterprise System/3090, Netview, PR/SM, Processor Resource/System Manager, Hiperspace, and ES/9000 are trademarks of the IBM Corporation. SAS, SAS/OR, and SAS/STAT are trademarks of the SAS Institute Inc. MXG is a trademark of Merrill Consultants. MICS is a trademark of Legent Corporation.

COPYRIGHT INFORMATION

Preface

Over the past several decades, computer technology advances have dramatically increased the processing power of central processors. Unfortunately, the improvements in I/O processing have not kept pace with the central processor advances. However, the data storage and accessing requirements of most installations continue to significantly increase as new applications are developed. The "DASD farm" often represents the largest investment in hardware in a large installation, and on-line applications require rapid access to data in order to meet response goals.

A variety of techniques have been implemented in an attempt to minimize the amount of I/O processing or to speed the access to data (dataspaces, virtual storage, expanded storage, and cached DASD are examples of techniques implemented to reduce the I/O requirement or to speed access). These techniques have had limited success in reducing the "I/O bottleneck" for most installations. Consequently, the DASD configuration still presents the largest potential for overall performance improvement and for performance improvement of specific applications. In fact, IBM has stated that "*over 75% of the problems reported to the IBM Washington System Center can be traced to some kind of I/O contention. Channel loading, control unit or device contention, data set placement, paging configurations, and shared DASD are the major culprits...*"[1]

We believe that a typical Data Base Administrator can solve the more serious DASD performance problems with such information as:

- Which DASD devices offer the most potential for improved performance.
- Which DASD devices cause the most adverse effect on the performance of specific application systems (such as on-line applications).
- Which applications use these DASD devices and how much the applications use the DASD devices.
- Which applications interfere with on-line workload (such as TSO or CICS), and what is the effect of the interference.

The DASD Component of CPExpert is designed to identify DASD problems such as those listed above, and to provide the information necessary to solve the problems. Data Base Administrators can then focus on correcting the DASD problems identified.

[1]MVS Performance Notebook, IBM publication GC28-0886

How to use this manual

This document describes how to use the DASD Component of CPExpert to analyze major constraints to improved performance of your computer system. The manual is organized into six sections and one appendix.

Section 1 provides an introduction to the DASD Component. This section is organized into four chapters. Most of this section can be reviewed for general information.

- Chapter 1 provides a brief background of DASD performance problems and performance analysis.
- Chapter 2 provides an overview of the DASD Component of CPExpert.
- Chapter 3 describes the sources of data used by the DASD Component to analyze DASD performance.
- Chapter 4 describes the performance data bases which CPExpert can use to analyze performance.
- Chapter 5 briefly describes the types of analysis that the DASD Component performs.

Section 2 provides information on installing the DASD Component (please follow the instructions contained in the *CPExpert Installation Guide* if CPExpert has not been installed). The instructions in this chapter should be followed closely when installing the DASD Component, and may be required when you make changes to your I/O configuration. This section is organized into three chapters:

- Chapter 1 provides detailed instructions on how to install the modification to MXG necessary to collect information relating DASD use to specific applications (jobs, job steps, and service classes or performance groups). If you use MXG to create your performance data base, the instructions in this chapter will optionally be followed when you initially install the DASD Component.
- Chapter 2 provides detailed instructions on how to install the modification to NeuMICS necessary to collect information relating DASD use to specific applications (jobs, job steps, and service classes or performance groups). If you use NeuMICS to create your performance data base, the instructions in this chapter will optionally be followed when you initially install the DASD Component.
- Chapter 3 provides detailed instructions on how to (1) define workload categories, (2) associate service classes or performance groups with the workload categories, and (3) direct CPExpert to perform analysis based upon the defined workload

categories. The workload definition is optional; you do not have define any workloads. However, you must define workloads if you wish CPExpert to analyze the DASD performance effects of contending workloads.

Section 3 describes how to provide guidance variables to the DASD Component using the CPEXPERT.SOURCE(DASGUIDE) PDS member. The instructions in this section will be important any time the guidance variables need to be changed. This section is organized into two chapters.

- Chapter 1 describes how to specify data selection and presentation guidance variables. The instructions in this chapter will be important if you wish to select specific measurement periods for analysis.

- Chapter 2 describes how to specify analysis control variables to guide the DASD Component in its analysis of system performance. The instructions in this chapter will be important if you wish to alter the defaults provided with the DASD Component.

- Chapter 3 describes how to exclude specific volumes from analysis.

- Chapter 4 describes how to select specific volumes for analysis.

- Chapter 5 describes how to analyze how well critical data sets meet response objectives.

- Chapter 6 describes how to analyze potential performance problems with VSAM data sets.

- Chapter 7 describes how to exclude specific VSAM data sets from analysis

Section 4 provides information on executing the DASD Component. The instructions in this section should be followed closely when executing the DASD Component. The instructions in this section will be important each time the DASD Component is executed.

Section 5 describes DASD performance considerations. The information in this section should be appreciated before attempting to use the DASD Component. The section is organized into two chapters.

- Chapter 1 provides an overview of DASD performance considerations. This chapter describes the major components involved in a typical DASD I/O operation. The information in this chapter allows you to appreciate the importance of the areas analyzed by the DASD Component.

- Chapter 2 describes limitations and considerations involved in using SMF/RMF data to analyze DASD performance. The information in this chapter is essential in determining whether the analysis performed by the DASD Component is appropriate for your environment.

Section 6 describes how to use the DASD Component to analyze performance. The instructions in this section should be followed each time you execute the DASD Component.

Appendix A contains a detailed description of each rule that results in a finding based upon the DASD Component analyzing performance of your DASD. You may wish to briefly review the rules in this appendix to appreciate the problems that are encountered in different installations. However, it is not necessary to read all of the rules. It is necessary only to read the rules that are identified by the reports produced from the DASCPE Module.

Acknowledgments

Computer Management Sciences would like to acknowledge the following individuals. They played a significant role in the concept, design, development, testing, or documentation of CPExpert[2]:

John Ebner, Systemhouse

Alan Greenberg, Social Security Administration

Stan Meacham, MCI Corporation

Barry Merrill, Merrill Consultants

Philip Mugglestone, European Software Product Services, NV

Bryant Osborn, US Department of Transportation

John Peterson, MCI Corporation

Bernie Pierce, IBM Corporation

Fred Voth, JOSTENS Corporation

CPExpert would not exist as a product if it had not been for the personal inspiration, professional advice, technical knowledge, and encouragement from these individuals!

Additionally, Computer Management Sciences would like to acknowledge the following individuals. They played a significant role in the concept, design, development, and testing of the DASD Component of CPExpert:

V. Lee Conyers, US Department of Transportation

Carson Ho, UNISYS

Bryant Osborn, US Department of Transportation

[2]The affiliation of the individuals is shown as of the time CPExpert was developed. Some of the individuals are no longer associated with the organizations shown.

Contents

Page

Section 3: Specifying Guidance Variables

Page

Section 4: Executing the DASD Component

Page

Section 5: DASD Analysis Factors

Section 6: Using the DASD Component

Appendix A: Description of Rules

Page

Exhibits

CPExpert Release 13.1:

The main changes to the DASD Component for CPExpert Release 13. are to:

- A new feature allows users to specify target response times for specific data sets. The DASD Component will use information from TYPE42DS records to detect data sets that miss a specified response objective. The devices on which the data sets reside will then be analyzed to identify performance problems.

- Variables have been added so a user can specify that CPExpert should ignore devices that are below a minimum I/O response, below a minimum I/O rate, or below a minimum total I/O wait time; and to suppress listing RMF intervals for devices that have no problems in the interval.

- The CPExpert modification to MXG or neuMICS has been modified to revised the "MAXDASD variable value exceeded:" message. This message can be produced by the CPExpert code that extracts device information from SMF TYPE30 records (as either MXG or neuMICS processes the SMF data). The previous message indicated that the MAXDASD variable was too low and some devices were ignored by CPExpert. One implication of this message was that user tended to increase the MAXDASD value to very large numbers. Since this value controls array sizes in many modules of the DASD Component, a large value for MAXDASD caused unnecessary overhead. Wit h the new approach, the code lists the jobs that have excess devices. A user can decide whether these jobs are sufficiently important to warrant the increased execution time of the DASD Component.

- The logic that processes TYPE42DS has been revised to significantly increase execution efficiency of the DASD Component.

- Add the ability to select up to 10 systems individually for analysis. Until Release 131, a user had the options of analyzing data for all systems in the performance data based, analyzing data for a specific sysplex (in case the performance data base contained data for more than one sysplex), or analyzing data for a specific system in the performance data base. With Release 13.1, up to 10 systems can be individually selected for analysis.

- Enhance the options provided with the SAS Output Delivery System (ODS). With Release 13.1, users who exercise the SAS ODS feature for creating CPExpert output can optionally create the output as a PDF file, which can be emailed to other users. Additionally, users can optionally specify a STYLE feature for either HTML or PDF output, if they have a preferred STYLE for HTML or PDF output. The optional links that are available with the HTML have been revised; SAS at some user sites did not create the HTML output in the "standard" way, and the CPExpert code that inserted links into the HTML output did not work properly. I have revised the code to place the links into the output as the output is created, rather than attempting to place the links into the final HTML output created by SAS.

- CPExpert now specifies OPTIONS COMPRESS=N; to override any site specification for file compression. Experiments have shown that CPExpert code runs significantly faster (using much less CPU time) if compression has been turned off.

CPExpert Release 12.2:

The main changes to the DASD Component for CPExpert Release 12.2 are to:

- Enhance the DASD Component to analyze problems with VSAM data sets[1]. This analysis is a partial automation of the analysis and guidance given in IBM's *VSAM Demystified* Redbook, SG24-6105. The VSAM analysis is performed if a MXG performance data base exists, and if SMF Type 42 (Data Set Statistics) and SMF Type 64 files exist in the MXG performance data base. These are MXG files TYPE42DS and TYPE64, respectively. The following rules have been added to the DASD Component as a part of the VSAM analysis:

  DAS600: Excessive Control Area (CA) splits occurred

  DAS604: Excessive secondary extents were allocated

  DAS605: Excessive extents were used and secondary allocation unit was small

  DAS606: Primary or Secondary allocation unit was small

  DAS607: VSAM data set is close to maximum number of extents

  DAS610: Relatively small CI size was used for sequential processing

  DAS611: Relatively large CI size was used for direct processing

  DAS612: Relatively large CI size was used for mixed processing

  DAS620: The number of data buffers should be increased

  DAS621: The number of index buffers should be equal to index levels

  DAS622: The number of index buffers should be more than STRNO value

  DAS625: NSR was used, but a large percent of the access was direct

[1] Thanks to Glenn Bowman (Wakefern Food Corporation, NJ), Joan Kelley (IBM Poughkeepsie, NY), and John Cothran (IBM Dallas, TX) for providing VSAM test data.

DAS635: LSR was used, but a large percent of the access was sequential

- Options are provided to analyze VSAM data sets residing only on "poorly performing" devices, analyze all VSAM data sets, analyze only VSAM data sets (suppressing the normal "worst devices" analysis), or suppress analysis of VSAM data sets.

- An option is provided to suppress analysis of DASD altogether, but simply create the model of the I/O configuration, and associated device/controller/channel activity. This option can be useful for reporting I/O activity for capacity planning.

- A report optionally is produced when common analysis is not performed because data sources are not available (or CPExpert has not been advised that the data sources are available). For example, the report will alert you to missing application analysis if the modification to MXG or MICS has not been made so CPExpert has application data available.

- A report optionally is produced if the DASD report is excessively large, and suggestions are given about how to reduce the size of the report.

- The DASD Component is updated to support z/OS Version 1 Release 4.

CPExpert Release 12.1:

The main changes to the DASD Component for CPExpert Release 12.1 are to:

- Completely revise the Component to eliminate the requirement that users provide IOCP macros from which CPExpert built a model of the I/O configuration. IBM now provides sufficient information in standard SMF records so the DASD Component can automatically create I/O configurations using SMF Type 70, Type 73, Type 74, Type 75, and Type 78CF. With Release 12.1 of CPExpert, users will no longer need to provide IOCP macros to the DASD Component, since CPExpert can obtain the information automatically.

- Eliminate all documentation related to IOCP macros and other documentation for user-provided input related to the I/O configuration.

- Provide data set access information related to those data sets (managed by DFSMS) residing on volumes with poor performance.

- Include a detailed analysis of PEND time, when PEND time is a major cause of performance problems.

- Include the ability to analyze more than one sysplex in a single execution of the DASD Component

- Enhance the output to provide a sysplex view of DASD volumes with poor performance, regardless of whether a specific system is being analyzed.

- Update the Component to support z/OS Version 1 Release 3.

CPExpert Release 11.2:

The main changes to the DASD Component for CPExpert Release 11.2 are to:

- Update the Component to support z/OS Version 1 Release 2.

- Add support for the SAS Output Delivery System (ODS) feature, to enable optional web access of CPExpert reports (this new option was suggested by **Harald Seifert** of HUK, Coburg, Germany).

- Add optional links in CPExpert reports (if the SAS ODS feature is invoked), that link rule output to CPExpert documentation for the rules produced.

- Revise the entire DASD Component User Manual to correct administrative errors and to make to document easier to use.

CPExpert Release 11.1:

The main changes to the DASD Component for CPExpert Release 11.1 are to update the Component to support z/OS Version 1 Release 1.

CPExpert Release 10.2:

The main changes to the DASD Component for CPExpert Release 10.2 are to update the Component to support OS/390 Version 2 Release 10.

CPExpert Release 10.1:

The main changes to the DASD Component for CPExpert Release 10.1 are to:

- Update the Component to support OS/390 Version 2 Release 9.

- Revise the DASD Component User Manual to correct Section 2, Chapter 4 (Installing the modification for MICS).

- Revise the DASD Component User Manual to remove the requirement that the GENPARMS module be executed.

CPExpert Release 9.2:

The main changes to the DASD Component for CPExpert Release 9.2 are to update the Component to support OS/390 Version 2 Release 8.

CPExpert Release 9.1:

The main changes to the DASD Component for CPExpert Release 9.1 are to:

- Update the Component to support OS/390 Version 2 Release 7.
- Correct documentation to reflect existing logic to handle more than 4 paths to a device.
- Add a description of techniques that are used to generate an IOCP input data set, using the Hardware Configuration Definition (HCD) component of MVS.
- Add documentation of Rule DAS102 and Rule DAS202 to describe the rules that have been a standard part of the output (but had been left out of the User Manual).
- Revise the description of installing the modification for MXG if you use the MXG BUILDPDB or use SAS/ITSV to process SMF data.
- Revise output based on suggestions from users.
- Correct code based on errors reported by users.

CPExpert Release 8.2:

The main changes to the DASD Component for CPExpert Release 8.2 are to:

- Update the Component to support OS/390 Version 2 Release 6.
- Correct code based on errors reported by users.

CPExpert Release 8.1:

The main changes to the DASD Component for CPExpert Release 8.1 are:

- Added code to support Workload Definitions for MVS (Goal Mode), when exercising the “Loved One” analysis option. The effect of this change is to allow users to specify

Service Classes for Goal Mode processing, rather than specify Performance Groups as done in pre-Goal Mode operation.

- Modified the DASDMXG module to include Report Classes for MVS (Goal Mode). Additionally, added code in all relevant modules to support Report Class processing.

Section 1: Introduction

This section provides a brief overview of DASD performance evaluation, provides an overview of the DASD Component of CPExpert, describes the sources of data used by the DASD Component to analyze system performance, and describes the performance data bases that the DASD Component can use.

Chapter 1: Background

From a global view, DASD analysis can be viewed from two perspectives;

- How the DASD space is allocated (how much space is allocated versus how much space is free), how the allocated space is used (whether the allocated space is actually used by the files), and how frequently DASD files are referenced by applications (whether the files should be on-line or perhaps stored off-line). DASD performance analysis from this perspective typically relates to the cost of maintaining the on-line storage represented by the DASD devices.

- How the DASD configuration and files effect overall system performance or effect the performance of individual applications. DASD performance analysis from this perspective deals with the day-to-day operational performance implications of the DASD configuration and files.

There are numerous products which examine DASD performance from the first perspective mentioned above. CPExpert does not depend upon these products, since a basic philosophy of CPExpert is to use information normally available in a performance data base. Consequently, the DASD Component does not analyze DASD space allocation.

The DASD Component focuses on DASD performance from the second perspective: whether performance constraints exist in the basic DASD configuration and to what degree particular applications impact the performance of other, more important, applications.

Beginning with Release 12.2 of CPExpert, the DASD Component also analyzes problems with VSAM data sets. This analysis is a partial automation of the analysis and guidance given in IBM's *VSAM Demystified* Redbook, SG24-6105.

Very sound DASD analysis advice was given by Friesenborg[1]: "It seems that the main thing to do is to avoid the outrageous; after that the subject doesn't warrant much attention." The major objective of DASD analysis is to identify and solve the "outrageous" or serious problems. The DASD Component of CPExpert attempts to detect the serious problems, identify the causes of the problems, and suggest solutions to the problems.

[1]Friesenborg, S. E., IBM Publication GG22-9370

DASD performance constraints can be corrected by a variety of actions at different levels:

- At the file level (e.g., change file block size, move files to different location)
- At the device level (e.g. upgrade device)
- At the controller level (e.g., change the number of paths, upgrade controller, shorten string, add cache)
- At the system level (e.g., reconfigure paths)
- At the operations level (e.g., schedule conflicting applications to different times)
- At the applications level (e.g., change storage accessing requirements or patterns).

The DASD Component provides information and recommendations regarding which actions are suitable for specific problems in your environment.

Chapter 2: The DASD Component of CPExpert

The DASD Component was implemented with the basic philosophy that:

- Most Data Base Administrators are able to address a limited number of DASD problems each day. The DASD Component was designed to identify the most serious DASD problems (both from the perspective of the overall DASD configuration and from the perspective of critical applications).

- Data Base Administrators are most interested in solving recurring DASD problems. The DASD Component was designed to identify the devices, files, and applications responsible for or experiencing recurring DASD problems.

With this design philosophy, the DASD Component (1) analyzes DASD information available in standard SMF/RMF records to select the most serious problems, and (2) identifies the devices and applications responsible for the problems or applications experiencing performance degradation because of the problems.

The guidance provided to the DASD Component allows a wide variety of options in the analysis process. These options allow users to define workload categories, to exclude specific DASD volumes from analysis, and to specify thresholds to limit the analysis.

These features and the analysis are provided by numerous SAS modules. These SAS modules (1) shape DASD performance and utilization data for detailed analysis by other modules, (2) determine the workload and analysis categories, (3) evaluate the data to assess potential areas of unacceptable performance, and (4) report the results from the evaluation. These modules are loaded and controlled by the central DASD Component of CPExpert (titled DASCPE).

- **Shape Data**. The majority of "processing" in the DASD Component is accomplished by numerous modules whose function is to combine, sort, summarize, correlate, and prepare the data for analysis. These modules extract and process relevant data elements from SMF records. Additionally, the modules incorporate a variety of queuing models to perform preliminary analysis of the data.

- **Determine workload and analysis categories**. A significant optional feature of the DASD Component is its ability to allow the specification of workload categories (e.g., TSO, CICS, Batch, etc.) and to focus analysis with respect to these workload categories. Additional controls provide the flexibility to exclude DASD volumes from consideration.

- **Evaluate Data**. The evaluation of the data is accomplished by rules whose purpose is to determine whether DASD performance was acceptable, and to analyze the DASD devices with unacceptable performance. These rules determine the devices with major problems, identify the probable causes of the problems, and (if expanded analysis is performed) identify the applications most likely responsible for the problems.

- **Report Results**. The DASD Component reports the results from the evaluation, in a variety of reports at the device or application workload level.

Each significant finding is described in Appendix A of this document. The description summarizes the finding, lists predecessor findings, discusses the rationale for the finding, and recommends action.

- The summary presents a short description of the finding.

- The predecessor findings are listed so you can follow the line of reasoning leading to a particular rule being executed.

- The discussion describes as much as necessary of the operation of the computer system as it relates to the particular finding. The purpose of the discussion is to explain the reasoning behind the finding. If appropriate, the discussion might refer you to related discussions in the DASD Component User Manual, or in a User Manual of another CPExpert component (e.g., the User Manual for the WLM Component may be used as a reference to avoid repeating a detailed discussion).

- The recommendations suggest possible actions that should be considered based on the findings. In many cases, multiple possible actions are listed. You must determine which actions should be taken (this determination is based upon the suitability of the actions to your own environment, the financial implications of the action, and the "political" acceptability of the action.)

Chapter 3: Data Sources

CPExpert analyzes the performance of the DASD configuration based upon data from three sources:

- **RMF information**. The basic analysis performed by the DASD Component relies on standard SMF Type 70(series) information. The DASD Component analyzes Type 70, Type 72, Type 74, Type 75, and Type 78CF information to build a model of the I/O configuration in the sysplex, and to analyze performance of devices that performed poorly relative to other devices in the configuration.

- **SMF information**. The DASD Component of CPExpert optionally includes a modification to MXG or to NeuMICS (using standard exit facilities). The modification allows extraction of SMF Type 30(Data Definition) information related to the individual DASD devices used by job steps . This information allows the DASD Component to perform "expanded" analysis, analyzing DASD performance based upon critical workloads.

  Additionally, the DASD Component optionally extracts data from SMF Type 42 (Data Set Statistics) to report information about the data sets residing on devices with poor performance.

- **Guidance information**. Guidance information is contained in CPEXPERT.USOURCE(DASGUIDE). This PDS member contains information that allows you to specify specific time periods in which CPExpert should analyze data, which performance groups (for Compatibility Mode) or service classes (for Goal Mode) are associated with different workload categories, what level of detail should be pursued in the DASD analysis, which volumes should be excluded from analysis, device path and caching characteristics, and other variables to control the operation of the DASD Component.

Chapter 4: Performance Data Bases

The "raw" RMF data contained in the SMF Type 30(DD) and Type 70(series) records must be translated into SAS format and placed into a performance data base before CPExpert can use the information. There are two ways in which these records can be placed into a performance data base:

- **MXG performance data base**. The performance data base can be created by Merrill's Expanded Guide (MXG) software. MXG is provided by Merrill Consultants, Dallas, Texas. MXG provides a low-cost mechanism by which installations can create and maintain a performance data base.

- **NeuMICS performance data base**. The performance data base can be created by the MVS Integrated Control System (NeuMICS). NeuMICS is provided by Computer Associates, Vienna, Virginia. The NeuMICS SMF (or Base) Component creates and maintains a comprehensive performance data base from SMF/RMF information.

The flexibility to use either of the above performance data bases is due to the fact that CPExpert is implemented in the SAS language. SAS provides a powerful macro facility, both with respect to macro coding and with respect to macro variable names.

CPExpert generally uses SAS macro variable names when referring to an element of information in the SMF records. CPExpert uses the SAS "%LET" statements to define the macro variables as MXG variable names or NeuMICS variable. These "%LET" statements are contained in unique variable definition modules for MXG and NeuMICS.

Thus, the same CPExpert software can be executed against either of the performance data bases, by only invoking the proper definition module. The SAS %LET statements in the definition module automatically cause CPExpert to refer to the proper MXG or NeuMICS data elements.

Chapter 5: Types of Analysis

The DASD Component performs two types of analysis of DASD performance and constraints to improved performance: **basic** analysis and **expanded** analysis. Additionally, the expanded analysis can analyze DASD performance from the perspective of specific application workloads or from the perspective of specific data sets. With CPExpert Release 12.2, the DASD Component analyzes potential performance problems with VSAM data sets. This section describes these types of analysis.

Chapter 5.1: Basic Analysis

The philosophy of the basic analysis is to identify the devices which have the most potential for improvement, and to analyze these devices in detail. The method used to identify a candidate set of devices is quite simple:

- CPExpert computes the average device response time for each **type** of device in the configuration, for each RMF measurement interval. The logic computes the average device response by type of device, since better performance would be expected from cached devices (for example) than from non-cached devices. This method essentially assesses the performance of each device against the performance of similar devices in the configuration.

- Devices that exceed the average device response time for their device type in any RMF measurement interval are selected as candidates for improvement. The rationale is that improvement efforts should not be directed at devices which provide better than average response. Thus, the candidate set of devices to analyze consists of those which provided worse than average response.

- The I/O rate of each candidate device is weighted by its response time. The result is a measure of the relative performance improvement **potential** of each device which provided worse than average response, from an overall system view. For example, consider two devices in a device type having an average I/O response of 20 milliseconds:

  Device A: I/O rate = 30 I/O operations per second
  Device response = 25 milliseconds
  Weighting factor = 30 * 25 = 750

  Device B: I/O rate = 5 I/O operations per second
  Device response = 40 milliseconds
  Weighting factor = 5 * 40 = 200

  In the above example, CPExpert would select Device A as having the most overall potential for improvement, even though its per-I/O device response was not as bad as the device response of Device B.

This logic totally ignores the potential situation where Device B holds critical data sets (an index, for example) and thus has more effect on overall system performance than the I/O rate indicates. Unfortunately, CPExpert is examining measurement data and has no way to assess relative importance of devices. (In the basic analysis, CPExpert **does** provide an option by which you can select specific devices, and CPExpert will analyze performance constraints of only those devices.)

- A model is built of the I/O configuration (using Type 70, Type 73, Type 74, Type 75, and Type 78CF information). Measurement data (channel, controller, device information) is processed against this model. CPExpert uses queuing models to compute the likely delays at various physical or logical parts of the configuration (for example, path delays, missed RPS reconnect delays, etc.). The result from this analysis associates probable delays with the specific devices.

- The device(s) with the overall most potential for improvement is selected for detailed analysis as to likely causes (seeking, missed RPS reconnect, etc.).

- The results from the above analysis are reported, normally in a one or two page report.

Chapter 5.2: Expanded Analysis (Specific applications)

CPExpert can perform an expanded analysis of DASD performance from the perspective of specific applications. This section describes how this analysis is accomplished.

- A modification is made to the MXG or NeuMICS software creating your performance data base. This modification normally uses standard user exits[2], adds an insignificant amount of processing time, and uses a small amount of DASD space.

  - The modification normally is invoked as MXG or NeuMICS process the SMF Type 30(Interval) records.

    For MXG, the modification **could** be invoked as MXG processes the SMF Type 30(Job Step Termination) records if you do not collect SMF Type 30(Interval) records. Using Job Step Termination records is discouraged since incomplete analysis may result (the job steps may span many RMF measurement intervals and it may be impossible to correlate the data recorded by RMF with lengthy job steps).

  - The modification records the system, job, job step, performance group or service class, and summary information about each job step's use of DASD devices by device number. The information is recorded only at the device level (rather than at the DD name level) and only required information is retained. Consequently, the records are small and require a small amount of DASD space.

- You define specific applications (or "loved one" applications) to CPExpert. Specific workload categories are defined (for example, TSO or CICS), and you define the performance groups or service classes assigned to the workload categories.

- CPExpert processes the data acquired using the modification to MXG or NeuMICS described above, and matches application use of devices against SMF Type 74 device information. The resulting information is application-specific, in that it relates device use to the applications using the devices.

- CPExpert then identifies the devices used by "loved one" workloads, computes the average device response from this set of devices, and selects all devices exceeding the average response as candidates for analysis.

- The normal basic analysis is performed of all devices resulting from the above selection criteria. The result is an analysis from the perspective of specific applications.

---

[2]A single one-line modification is required to NeuMICS code in the NeuMICS DYSMFFMT module. NeuMICS does not currently provide an exit which can be used to add a data set to the output from NeuMICS as it analyzes SMF data. Such an exit does exist for the NeuMICS CICS Component (the _USRSDKP exit), and a request has been made that a similar exit be provided for the NeuMICS BASE Component.

- Additionally, CPExpert can identify non-critical applications causing performance problems for the "loved one" workload. For example, CPExpert can identify all non-critical applications using a significant amount of a path if CPExpert discovers that path contention results in poor performance to "loved one" applications.

Chapter 5.3: Expanded Analysis (Specific data sets)

CPExpert can perform an expanded analysis of DASD performance from the perspective of specific data sets. The analysis is done in one of two ways: (1) analyzing data set response based on information from TYPE42DS statistics or (2) analyzing data based on information in TYPE14/15 and using the CPExpert modification to MXG or NeuMICS as the SMF Type30 records are processed.

- **Analysis based on TYPE42DS.** This method can provide comprehensive information, without requiring the modification to MXG. This method **is not applicable** to NeuMICS, since NeuMICS does not retain sufficient information related to SMF Type 42 (Data Set Statistics).

  MXG creates TYPE42DS from the SMF Type 42 (Subtype 6 - data set I/O statistics) records created by SMS. The TYPE42DS file contains I/O access characteristics information, at the data set level. CPExpert extracts data set information and data set response statistics from TYPE42DS, and compares these to the response objectives you have made in USOURCE(DASGUIDE) about critical data sets. When any data set response time exceeds the specified objective, the DASD Component selects the data set and the volume it resides on for detailed analysis.

- **Analysis based on TYPE14/15 and CPExpert modification**. This method is a bit more involved, but can be used regardless of whether data sets are managed by SMS, and this method can be used regardless of whether you use MXG or NeuMICS to create your performance data base.

  With this method, the DASD Component analyzes SMF Type 14/15 records to extract data set names that correspond to the critical data sets that you have identified in USOURCE(DASGUIDE). The CPExpert modification to MXG or NeuMICS extracts DD information as SMF Type 30 records are processed. The DASD Component then correlates the data set information with the DD information, to determine whether critical data sets exceed the specified response objective. When any data set responses time exceeds the specified objective, the DASD Component selects the data set and the volume it resides on for detailed analysis.

With either of the above methods, CPExpert matches the data set information against SMF Type 74 device information, and identifies each device which has an average I/O response time exceeding the response objective of any data set residing on the device. The resulting information is specific to devices which have I/O response times exceeding a response objective.

The I/O rate of each device identified above is weighted by its response time. The result is a measure of the relative performance improvement potential of each device which provided I/O response exceeding the response objective for some data set on the device.

The normal basic analysis is performed of all devices resulting from the above selection criteria. The result is an analysis from the perspective of specific data sets.

Additionally, CPExpert creates a report describing the I/O activity of all data sets which resided on devices having an I/O response worse than the response objective for the data set.

Chapter 5.4: Analysis of shared DASD conflicts

CPExpert can perform an analysis of conflicts between DASD shared by systems or MVS images. This section describes how this analysis is accomplished.

Assume that CPExpert is analyzing DASD performance problems with System A, and some DASD devices are shared with System B. There are three major ways in which the shared DASD can cause performance problems for System A:

- **Case #1**. The arm of the device has been moved by System B when System A attempts to access a cylinder. The DASD I/O operation from System A must move the arm to the desired cylinder, as a SEEK operation.

  CPExpert computes the time required to perform SEEKs on System A. If the computed SEEK time is a major cause of performance problems, CPExpert analyzes the data from System B to determine whether System B generates a large number of I/O operations to the device.

  - If System B does **not** generate a relatively large number of I/O operations to the device, CPExpert concludes that there is **not** a conflict. There is little doubt about the validity of this conclusion. If System B does not direct much I/O activity to the device, System B clearly cannot cause seek problems for System A.

  - If System B **does** generate a relatively large number of I/O operations to the device, CPExpert concludes that there **is** a conflict caused by sharing the device.

    To be absolutely correct, CPExpert should process the configuration definitions for System B, process System B's channel and device information, and compute seek information for System B. CPExpert could then determine whether System B also experienced a high seek rate for the device.

    If both System A and System B experienced a high seek rate, CPExpert could be absolutely sure that there was a shared DASD conflict. This approach would unnecessarily use system resources and would be cumbersome to implement.

    Consequently, CPExpert makes the assumption that I/O operations to the device are random between System A and System B. CPExpert thus can conclude that if System A experiences a high seek rate and System B significantly uses the device (exhibited by a high I/O rate), then System B must also experience a high seek rate. To assume otherwise would require that I/O from System B be coordinated with the I/O from System A, such that System B does not experience seeking similar to System A.

    Since System A experiences a high seek rate and System B experiences a high I/O rate, CPExpert concludes that there is a conflict caused by the shared DASD.

- **Case #2**. The device (or the path to the device) is busy to System B when System A attempts to reference the device. The DASD I/O operation from System A must wait until the device (or path) is free. This wait time is directly reflected as PEND time in the RMF measurement data.

  CPExpert determines whether PEND time is a major cause of performance problems for the device (or devices) selected for detailed analysis. If PEND time is a major cause of performance problems, CPExpert analyzes the data from System B to determine whether DASD I/O operations from System B could cause the PEND delays for System A.

  CPExpert analyzes the total CONN time and DISC time for the device from other systems sharing the device. CPExpert concludes that the PEND delay is caused by other systems if the total CONN time and DISC time is more than 25% of the PEND delay for the device experienced on System A.

- **Case #3**. The device is cached, the cached controller is shared between System A and System B, and I/O operations from System B preempt data in the cache which is required by System A.

  This problem is evidenced by relatively large I/O times for System A. Since the data required by System A is not in the cache, the data must be acquired from the device. This potentially results in a seek operation to position the arm, and will require normal latency and data transfer from the device.

  In any of the above situations with cached shared devices, CPExpert analyzes the amount of I/O operations from System A and System B to the cached controller.

  - If System B does **not** generate a relatively large number of I/O operations to the device, CPExpert concludes that there is **not** a conflict.

  - If System B **does** generate a relatively large number of I/O operations to the device, CPExpert concludes that there **is** a conflict caused by sharing the device. The rationale for these conclusions is the same as was discussed under Case #2 (seek conflicts).

  The standard analysis performed by CPExpert may not detect a cache problem under two possible scenarios.

  - If only one cached controller is attached to System A, CPExpert may not detect a problem with the device. This is because the logic employed by CPExpert selects devices with the most performance improvement within each **type of device** and then selects the **overall** "worst" devices for detailed analysis.

    CPExpert considers cached devices to be a unique type. If all devices on the cached controller received bad service caused by shared cached problems, CPExpert may not detect a performance problem with the cached devices. This is because all devices in the "device type" could have roughly equal poor service and no device would be **significantly** worse than the other devices in the device type. Consequently, CPExpert might not select any of the cached devices for detailed analysis

If there are multiple cached controllers, there would be a larger number of "candidate" devices, and the standard analysis performed by CPExpert is more likely to identify any problem caused by shared cache controllers.

- The analysis performed by CPExpert may not detect a cache problem if the cache is being replaced with data from another volume. It is possible that System B could cause data from another volume to be loaded into the controller's cache. This volume could be a volume not be flagged as the "worst" performing volume when analyzing performance from the perspective of System A. From System A's perspective, accesses to the worst volume would simply not find required data in cache.

  CPExpert cannot at present relate (1) poor performance for one volume on System A caused by cache problems and (2) I/O operations to a different volume by System B.

The shared DASD analysis is accomplished when you perform the following steps:

- You collect and process SMF data for each system to be analyzed. The SMF data must be placed into your standard performance data base. This step is not specific for CPExpert; it is the normal processing you would do. The step is listed first just for completeness.

- **You should ensure that the system clocks on all systems are set at roughly the same time**. CPExpert will examine the SMF data contained in your performance data base, based upon the SMF time associated with the measurement data. Shared DASD problems between systems will be detected based on this SMF time. If the SMF times are significantly different, the analysis might not properly identify conflicts or conflicting applications.

  You should not worry that the system clocks be set at **exactly** the same time. The analysis performed by CPExpert is not intended to identify an isolated performance problem. Rather, CPExpert attempts to identify those problems which **continually** cause shared DASD performance problems. If a shared DASD problem is continual, the problem will be reflected in multiple SMF recording intervals. Consequently, it is not essential that the SMF times be exactly matched between systems.

- You specify **%LET SHARED = YES;** in USOURCE(DASGUIDE) to tell CPExpert to perform analysis of shared DASD.

CPExpert performs the following processing if you have indicated that an analysis of potential conflicts between shared DASD should be performed:

- CPExpert determines whether the "worst" device selected for detailed analysis is attached to a control unit shared with another system. If so, CPExpert performs an analysis of potential conflicts caused by shared DASD.

- CPExpert identifies other systems which reference the "worst" device by analyzing the SMF Type 74 data in the performance data base relating to all other systems. The SMF Type 74 data contain the VOLSER for each device referenced. CPExpert simply selects SMF Type 74 information for the systems which reference the VOLSER of the "worst" device. This information is retained for more detailed analysis about potential conflicts.

- Once CPExpert has identified all systems which reference the "worst" device, CPExpert analyzes the DASD I/O characteristics of these systems with respect to the "worst" device. As described earlier, the analysis makes a basic assumption that the I/O activity from the different systems is random among the systems (that is, the I/O activity of System B is independent from the I/O activity of System A).

  Consequently, CPExpert can conclude that delays on System A are caused by shared DASD conflicts only if (1) System B experienced the same delays with the "worst" device (in the case of PEND delays) or (2) System B generated a high I/O activity to the "worst" device. CPExpert can conclude that delays on System A are not caused by shared DASD conflicts if the data from System B does not exhibit these characteristics. The earlier discussion explains the rationale for these conclusions.

- If "expanded" analysis is being performed, CPExpert can identify the applications on System B which reference the "worst" device.

  SMF Type 30 records do not contain the VOLSER in the DD section. However, CPExpert correlates the information acquired from other records to relate the VOLSER to the device number. While extracting data from the SMF Type 74 information in the performance data base, CPExpert retains the device number associated with the devices. This device number is matched against the device numbers in the SMF Type 30 records to identify the VOLSER associated with the device.

  CPExpert thus can identify all applications from System B which reference the "worst" device shared with System A.

Chapter 5.5: Analysis of VSAM data sets

CPExpert can perform an analysis of problems with VSAM data sets, **if MXG has created the performance data base**. This section describes how this analysis is accomplished.

VSAM data set activity typically accounts for a large percent of I/O activity (more than 70% at some sites). Tuning of a few files or correcting common problems often can result in significantly improved performance (IBM benchmarks show up to 90% improvement resulting from some simple changes).

With CPExpert Release 12.2, the DASD Component has been enhanced to provide a rudimentary analysis of common VSAM problems. Additional analysis will be added in future enhancements to the DASD Component.

The DASD Component can optionally analyze VSAM data set performance problems or potential problems under the following conditions:

- MXG has created the performance data base. Unfortunately, there is insufficient data retained in a NeuMICS performance data base to allow CPExpert to analyze VSAM performance problems.

- The performance data base contains the MXG TYPE64 file (and CPExpert is provided with the **%LET TYPE64=Y;** guidance variable).

- Most analysis of VSAM data sets also requires the MXG TYPE64 file (and CPExpert is provided with the **%LET TYPE64=Y;** guidance variable). This is because the analysis depends on being able to identify the VSAM file type (e.g., KSDS, VRRDS, etc.), and the buffering technique used (e.g., NSR or LSR).

Section 2: Installing the DASD Component

Most of the DASD Component will be installed as normal part of installing CPExpert. However, up to three additional steps may be taken: (1) installing a modification to MXG or to NeuMICS to collect application-related or data set-related information, (2) defining workload categories, and (3) specifying data set information.

- **Collecting application-related or data set-related information**. This is an optional step. The DASD Component can optionally identify the specific applications (jobs and job steps) which access specific DASD volumes or use specific controllers or channel paths, or can analyze devices which provide a device I/O response time exceeding the response objective for selected data sets. This information can be quite useful in solving a DASD performance problem. With the application-related analysis, the DASD Component identifies the applications causing the problem or can identify the applications suffering from the problem. With the data set-related analysis, the DASD Component identifies devices which provide an I/O response time that exceeds a response objective for any critical data sets residing on the device. Providing this information is possible only if a modification is made to MXG or to MICS, to alter their processing of SMF Type 30(Data Definition) information.

  The modification included with CPExpert is quite simple.

  - For the application-related analysis, CPExpert records only the essential information required to identify the job name, the step name, the performance group number (for Compatibility Mode) or service class (for Goal Mode), the devices referenced and their SIO and connect times, and the beginning and end time of the measurement. This information is sufficiently concise that less than 25 cylinders of DASD are required to hold the information for all job steps executed in a relatively large installation. However, the information is sufficiently comprehensive for CPExpert to perform substantial analysis of DASD performance and the causes of poor DASD performance.

  - For the data set-related analysis, CPExpert records information associated with specific DD names: the job name, the step name, the performance group number or service class, the device referenced and the SIO and connect time for the device, and the beginning and end time of the measurement.

  Section 1 (Chapter 5) of this user manual describes the modifications and the processing of SMF information that CPExpert performs.

  Chapter 1 describes how to install the modification for MXG code, and Chapter 2 describes how to install the modification for MICS code. The installation for either product is very straightforward.

- **Defining workload categories**. This is an optional step. CPExpert allows you to define categories of work, and have the analysis focus on specific workload categories.

This can be useful, for example, if you wish to identify your "loved ones" and have the analysis treat this workload category as more important than other work.

Chapter 3 of this section describes how to define workload categories to CPExpert. |

- **Specifying data set names**. This is an optional step. CPExpert allows you to specify data set names and associate a response objective with each data set. CPExpert will analyze SMF data to detect periods when devices on which the data sets reside provide an I/O response greater than the response objective. If any such intervals are detected, CPExpert will perform a detailed analysis of the devices to identify why the device I/O response exceeded the response objective.

  Chapter 4 of this section describes how to define data set names and response objectives to CPExpert. |

The DASD Component provides many processing options (for example, you may exclude certain volumes from analysis, select particular time intervals for analysis, etc.). These processing options are not, strictly speaking, a part of the installation process. They may be employed as you refine your analysis of the performance of your DASD configuration. These processing options are described in Section 3 of this User Manual.

The instructions in this section assume that you have installed the CPExpert software. If you have not installed CPExpert, please install the software before continuing. (The procedures for installing CPExpert are described in the *CPExpert Installation Guide*.)

Chapter 1: Installing the modification for MXG

This chapter applies only if you use MXG to maintain your performance data base.

With normal MXG processing, the MXG TYPE30_D data set is not written because the data set would generally require **far** too much DASD space (often an entire volume would be required to hold the data set). Consequently, Barry Merrill does not write the TYPE30_D data set and he includes appropriate warnings in the IMAC30DD module that the data set may be very large. The CPExpert modification to MXG can allow you to collect essential information relating the DASD use to specific job steps and performance groups or service classes (with Goal Mode), without requiring massive amounts of DASD space. Additionally, the modification can allow you to collect information required to relate I/O activity to specific data sets.

Installing the modification for MXG is an optional step; however the DASD Component will be unable to relate DASD use and performance constraints to specific applications or to specific data sets unless you install the modification.

Implementing the DASD Component modification to MXG consists of five steps:

Step 1: Install the DASDMXG code.

Install the DASDMXG code into the SAS source library containing user modifications to MXG. This is done by copying the DASDMXG member from the CPEXPERT.SOURCE PDS into MXG.USERLIB (or whatever you call your library for user modifications to MXG). The DASDMXG member contains all the macros that will be invoked by the modifications described below.

Step 2: Modify MXG modules.

Modify one of two MXG modules to cause the DASD Component to output the DASD30DD data set. The MXG modules that may be modified are the IMACINTV module or the EXTY30U4 module.

- If you collect SMF Type 30 Interval records, modify the MXG module IMACINTV by including the following statement in the module:

```
%CPEOUTDD(V);      /* CPEXPERT MODIFICATION */
```

Exhibit 2-1 illustrates the MXG IMACINTV module with the modification.

```
/*  COPYRIGHT (C) 1985,1992 BY MERRILL CONSULTANTS DALLAS TEXAS     */
/*                                                                  */
/* IMACINTV CONTROLS THE CREATION OF TYPE30_V  DATASET, BUILT FROM  */
/* TYPE 30 INTERVAL RECORDS (SUBTYPES 2 AND 3).                     */
/*  NOTE THAT THE PDB.SMFINTRV DATASET IS A RENAME OF TYPE30_V AND  */
/*  THUS CHANGING THIS MEMBER TO CREATE TYPE30_V OBSERVATIONS WILL  */
/*  ALSO CREATE PDB.SMFINTRV IF YOU EXECUTE BUILDPDB/BUILDPD3.      */
/* BY DEFAULT, THIS MEMBER CREATES NO OBSERVATIONS IN TYPE30_V.     */
/* IF YOU WANT ANY OBSERVATIONS TO BE CREATED IN THE DATASET, YOU   */
/* MUST EXECUTE AN   OUTPUT _LTY30UV;  STATEMENT IN THIS MEMBER.    */
/* FOR  EXAMPLE, IF YOU WANT AN OVSERVATION IN TYPE30_V FOR EACH AND*/
/* EVERY TYPE 30 SUBTYPE 2 OR SUBTYPE 3 RECORD READ, SIMPLY REMOVE  */
/* THE COMMENT BLOCK AROUND THE FOLLOWING STATEMENT:                */
/* FOLLOWING STATEMENT:                                             */
/*                                                                  */
/*        OUTPUT _LTY30UV ;                                         */
/*                                                                  */
/* THE _LTY30UV MACRO, DEFINED IN IMAC30, DEFAULTS TO TYPE30_V.     */
/*                                                                  */
/* ALTERNATIVELY, IF YOU  WISH  TO  CREATE  OBSERVATIONS  ONLY  FOR */
/* STARTED  TASKS,  REMOVE  THE  COMMENT  BLOCK  FROM THE FOLLOWING */
/* STATEMENT:                                                       */
/*                                                                  */
/*        IF TYPETASK='STC' THEN OUTPUT _LTY30UV;                   */
/*                                                                  */
/* IMACINTV IS %INCLUDED AFTER THE "IMPORTANT" VARIABLES, SUCH AS   */
/* JOB, READTIME, STEPNAME, PROGRAM, TYPETASK, JOBCLASS, SMFTIME,   */
/* ETC. HAVE BEEN DEFINED.  THUS YOUR SELECTION CRITERIA CAN CAUSE  */
/* TYPE30_V TO CONTAIN OBSERVATIONS ONLY  FROM SPECIFIC JOBS AND/OR */
/* PARTICULAR PROGRAM NAMES, SUCH AS                                */
/*                                                                  */
/*   IF JOB=:'CICS' OR PROGRAM='DFHSIP' THEN OUTPUT _LTY30UV;       */
/*                                                                  */
/* REALIZE THAT THIS MEMBER CONTROLS ONLY THE SELECTION OF TYPE 30  */
/* INTERVAL RECORDS. YOU MUST (IN PARMLIB) SPECIFY THAT YOU DESIRE  */
/* INTERVAL RECORDS TO BE CREATED BY SMF; YOU CAN SPECIFY SEPARATE  */
/* DURATIONS INDIVIDUALLY FOR JOBS (JOB), TSO SESSIONS (TSU), AND   */
/* OR STARTED TASKS (STC).                                          */
/*                                                                  */
/*                                                                  */
/*                                                                  */
/*                                                                  */
/******************************************************************/
/***   THIS MODULE IS INCLUDED BY THE FOLLOWING MXG MODULES:      ***/
/***                                                              ***/
/***     VMAC30                                                   ***/
/***                                                              ***/
/******************************************************************/

%CPEOUTDD(V);              /* CPEXPERT MODIFICATION     */
```

MXG IMACINTV MODULE, WITH CPEXPERT MODIFICATION

EXHIBIT 2-1

MXG coding printed with permission of Merrill Consultants, Dallas, Texas.

The macro parameter "(V)" in the above code causes SAS statements to be generated to identify the source of the data (namely, the SMF Type 30 Interval records).

Collecting interval data is BY FAR the preferred approach, since much more comprehensive analysis can be accomplished.

- If you do **not** collect interval data but collect only Workload Termination data (that is, you collect SMF Type 30 Workload Termination records but do not collect Type 30 Interval records), you can still obtain information about the DASD use of specific job steps. However the analysis will not be as comprehensive as if you collect interval information.

  If you collect only Workload Termination records, modify the MXG module EXTY30U4 by including the following statement in the module:

```
%CPEOUTDD(4);    /* CPEXPERT MODIFICATION     */
```

The macro parameter "(4)" in the above code causes SAS statements to be generated to identify the source of the data (namely, the SMF Type 30 Workload Termination records.

Exhibit 2-2 illustrates the MXG EXTY30U4 module with the modification.

```
/* COPYRIGHT (C) 1985,1992 BY MERRILL CONSULTANTS DALLAS TEXAS USA */
/****************************************************************/
/*  MEMBER  EXTY30U4  -  MXG  OUTPUT EXIT FOR DATA SET TYPE30_4  */
/****************************************************************/
/*  THIS MEMBER CONTAINS THE OUTPUT STATEMENT FOR THE DATA SET.  */
/*  CHANGES IN DDNAME, DSNAME, AND THE CREATION OF NEW VARIABLES */
/*   CAN BE CODED HEREIN.                                        */
/****************************************************************/
  OUTPUT _LTY30U4; /* TYPE30_4, DEFINED IN IMAC30 */

 %CPEOUTDD(4);                   /* CPEXPERT MODIFICATION        */
```

MXG EXTY30U4 MODULE, WITH CPEXPERT MODIFICATION

EXHIBIT 2-2

MXG coding printed with permission of Merrill Consultants, Dallas, Texas.

Step 3: Modify MXG IMAC30DD module.

Modify the MXG IMAC30DD module to cause the DASD Component to collect the Data Definition information that is contained in the DASD30DD data set. This is accomplished by including the following statement in the module:

```
%DASD_MXG;    /* CPEXPERT MODIFICATION    */
```

Exhibit 2-3 illustrates the MXG IMAC30DD module with the modification.

```
/* COPYRIGHT (C) 1985,1992 BY MERRILL CONSULTANTS DALLAS TEXAS */
/**********MEMBER=IMAC30DD**************************************/
/*                                                             */
/* THIS MODULE CONTROLS THE EXISTENCE OF DATA SET   TYPE30_D   */
/*  (WHICH IS RENAMED PDB.DDSTATS BY BUILDPDB/BUILDPDE).       */
/*                                                             */
/* BECAUSE TYPE30_D COULD CONTAIN ONE OBSERVATION FOR EVERY DD */
/*  CARD IN EVERY SUBTYPE OF EVERY TYPE 30 RECORD, THE DATASET */
/*  COULD BE VERY, VERY LARGE, AND THUS BY DEFAULT, MXG DOES NOT */
/*  CREATE ANY OBSERVATIONS IN TYPE30_D.                       */
/*                                                             */
/* YOU MUST MODIFY THIS INSTALLATION MACRO TO CREATE OBSERVATIONS! */
/*                                                             */
/* NOTE THAT DATASET TYPE30_D AND PDB.DDSTATS WILL ALWAYS EXIST; */
/*  THIS EXIT CONTROLS ONLY IF THERE ARE OBSERVATIONS CREATED. */
/*                                                             */
/* THE DSNAME MACRO _LTY30UD IS DEFINED IN IMAC30 AND DEFAULTS */
/*  TO TYPE30_D.                                               */
/*                                                             */
/* THE FOLLOWING EXAMPLE HAS BEEN COMMENTED OUT SO THAT IT IS NOT */
/*  EXECUTED, AND IT SHOWS ONE OF MAY CONDITIONS YOU MIGHT CHOOSE */
/*  TO RESTRICT WHICH DD SEGMENT'S OBSERVATIONS ARE OUTPUT.  THE */
/*  EXAMPLE WILL CREATE OBSERVATIONS ONLY FOR THE STEP TERMINATION */
/*  (SUBTYPE=4) RECORD, AND ONLY FOR THE SPECIFIC DEVICE NUMBER */
/*  (UCB ADDRESS=38F) IN THIS EXAMPLE.  YOU COULD CHOSE TO TEST */
/*  FOR JOB, STEP, DDNAME, OR ANY OF THE VARIABLES IN TYPE30.  */
/*                                                             */
/* NOTE: THE EXAMPLE TEST IS INSIDE A PAIR OF COMMMENT SYMBOLS */
/*       WHICH PRECEED AND FOLLOW THE EXAMPLE "IF STATEMENT".  */
/*       YOU MUST REMOVE THE COMMENT SYMBOLS AROUND THE TEST, AND */
/*       CHANGE THE TEST AS DESIRED, TO CREATE OBSERVATIONS.   */
     /*
         IF DEVNR=038FX AND SUBTYPE=4 THEN OUTPUT _LTY30UD;
     */

/*******************************************************************/
/***   THIS MODULE IS INCLUDED BY THE FOLLOWING MXG MODULES:     ***/
/***                                                           ***/
/***     VMAC30                                                ***/
/***                                                           ***/
/*******************************************************************/
/*                                                                 */

%DASD_MXG;  * CPEXPERT MODIFICATION TO ACQUIRE DASD INFORMATION;
```

MXG IMAC30DD MODULE, WITH CPEXPERT MODIFICATION

EXHIBIT 2-3

MXG coding printed with permission of Merrill Consultants, Dallas, Texas.

Step 4(alternate): Modify MXG EXPDBINC module and EXPDBVAR module.

This step applies if you use the standard MXG BUILDPDB or use SAS/ITSV to process SMF data.

Modify the MXG EXPDBINC module as shown below:

```
%INCLUDE SOURCLIB(DASDMXG);    /* CPExpert modification */
```

Modify the MXG EXPDBVAR module as shown below:

```
MACRO _VARUSER      /* CPExpert modification */
    %%VARDDCPE(MAXDASD=nn,MAXDD=nnn,MINIO=nnn,COLLECT=xxxxxxx)
%
```

The %%VARDDCPE macro statement generates code to define the DASD30DD data set, define the DASD30DS data set, and describe the variables contained in the KEEP statement for the data sets. (**NOTE**: two "%" symbols are required since the code is placed into SAS "old style" macros.) The macro parameters specify the maximum number of unique DASD devices referenced by any job step in your environment, the maximum number of DD statements associated with critical data sets, the significant number of I/O operations directed to any data set in any Type 30(Interval) record, and which type of data collection should be performed.

- The **MAXDASD** macro parameter specifies the maximum number of unique DASD devices that are referenced by any job step in your environment. Note that this is **not** the number of DD statements. Rather, this is the number of unique devices. There often will be many DD statements for a single device, but the DASD_MXG software combines these DD statements into one observation in the DASD30DD data set.

  The default specification for the MAXDASD macro parameter is **MAXDASD=25**. The DASD_MXG software will produce a warning message if this value is too low, and the software simply ignores devices for any job step exceeding the value specified. (There is no other effect; that is, the software will not abort and the analysis software will continue to function properly.)

  The only reason for specifying 25 versus 200 (for example) as an initial value, is that the larger the value, the more DASD space that will be used for the DASD30DD data set. The record size will increase by 15 bytes for each unique device specified. If you process 20,000 job steps each day and specify a maximum of 25 unique devices per job step, the total temporary DASD space required for the DASD30DD data set will be about 25 cylinders (3380). This is not a large amount in any event, but there is no point in writing larger records than are needed.

Therefore, MAXDASD=25 should be specified initially and the value should be increased only if required.

- The **MAXDD** macro parameter specifies the maximum number of unique DD names that may be associated with critical data sets. As described in Section 1 (Chapter 5.3), the CPExpert modification to MXG creates an array which contains the DD names of all DD statements associated with data sets which have been defined to CPExpert with a response objective. (The array is created only if you specify DDNAMES or BOTH in the COLLECT macro parameter described below.)

  Each array element requires only 8 bytes of storage. This amount storage is insignificant. However, SAS generates a SAS data element name for each array element. SAS 5.18 has a restriction on the total number of data element names which can be defined. When this number is exceeded, SAS aborts with an error in the compilation stage. Therefore, you cannot set an extremely large value for the MAXDD parameter when executing under SAS 5.18.

  The default specification for the MAXDD macro parameter is **MAXDD=200**, generating an array capable of containing 200 unique DD names. The DASD_MXG software will produce a warning message if the number of unique DD names is greater than the array size. The software will list all DD names which it could not process, and it simply ignores DD names for any job step exceeding the value specified. (There is no other effect; that is, the software will not abort and the analysis software will continue to function properly.)

  If you receive a warning message from the CPExpert modification, you should consider (1) increasing the value of the MAXDD macro parameter or (2) restricting the number of data sets you define as critical.

- The **MINIO** macro parameter specifies the minimum number of I/O operations to any data set within a particular SMF Type 30(Interval) record. The CPExpert modification will ignore any data sets encountered if the number of I/O operations is below the MINIO value. The point of this parameter is that you probably will not worry about data sets with a very low activity, regardless of the DASD response time they experienced. There is no point in collecting information about very low activity data sets.

  The default specification for the MINIO macro parameter is **MINIO=100**.

- The **COLLECT** macro parameter specifies whether to collect device information (required to associate device activity with application systems), to collect DD name information (required to associate device activity with data sets), or to collect both device and DD name information.

Specify **COLLECT=DEVICES** to collect device information. This is the default setting. If you have previously installed the modification for MXG, you do not have to make any changes to your installation.

Specify **COLLECT=DDNAMES** to collect DD name information.

Specify **COLLECT=BOTH** to collect both device information and DD name information.

WARNING: Please review Chapter 6 of this section before specifying DDNAMES or BOTH as an option for the COLLECT parameter. You should **not** specify either of these options unless you have defined critical data sets and have followed the procedures in Chapter 6.

Step 4(alternate): Modify the SAS job stream used to execute MXG.

This step applies if you use in-stream JCL for MXG to process SMF data. Modify the SAS job stream used to execute MXG by making the following simple modifications:

- Modify the %INCLUDE statement to add **DASDMXG** to the list of included code.

- Add the following macro statement after the MXG DATA statement:

%VARDDCPE(MAXDASD=nn,MAXDD=nnn,MINIO=nnn,COLLECT=xxxxxxx)

Please refer to the discussion in "Step 4 (if you use the standard MXG BUILDPDB to process SMF data)" for a description of the macro variables.

WARNING: Do not include a semicolon after the macro invocation!

Exhibit 2-4 illustrates a sample MXG SAS job stream with the three modifications highlighted. Of course, your own MXG SAS job stream may appear somewhat different than that shown in Exhibit 2-4. However, your MXG SAS job stream should be sufficiently similar to Exhibit 2-4 that you can appreciate how the modifications should appear in your own MXG SAS job stream.

Step 5: Add CPEDASD DD statement to the JCL.

Add the CPEDASD DD statement to the job control language (JCL) used to execute MXG. The CPEDASD DD statement defines the SAS library in which the CPExpert modification will place the DASD30DD data set or DASD30DS data set created by the modification.

The CPEDASD DD statement can refer to the CPEDATA SAS library if you wish, or you can allocate a different SAS library. For space allocation purposes, the CPEDASD data set requires slightly less than one cylinder of DASD space per 2,200 SMF Type 30(Interval) records processed by MXG. For example, if you typically have about 22,000 Type 30(Interval) records in your daily SMF file, you will need about 10 cylinders of DASD allocated to the CPEDASD SAS library. It is difficult to estimate the amount of space required for the DASD30DS data set, since this is a function of the number of critical data sets you define and the number of unique DD names associated with the data sets. This data set contains about 730 observations per track.

Exhibit 2-4 illustrates a sample MXG SAS job stream with the CPEDASD DD statement included, using the CPEDATA SAS library space.

```
//CPEDASD      DD      DSN=prefix.CPEXPERT.CPEDATA,DISP=OLD
//SYSIN        DD      *

%INCLUDE
 SOURCLIB(VMACSMF,VMAC7072,VMAC71,VMAC73,VMAC74,VMAC75,VMAC77,
          VMAC78,VMAC30,VMACTSOM,VMAC434,VMAC40 VMAC434D,
          DASDMXG);
DATA

  %VARDDCPE(MAXDASD=25,MAXDD=200,MINIO=100,COLLECT=DEVICES)

  _VAR7072
  _VAR71


etc.


;
```

SAMPLE MXG SAS JOB STREAM

EXHIBIT 2-4

 Revised: April, 2003

Chapter 2: Installing the modification for NeuMICS

This chapter applies only if you use MICS to maintain your performance data base.

The normal MICS processing summarizes the device information from the SMF Type 30(DD) EXCP segment into the BATWDA dataset. The BATWDA dataset is created only if specified during the MICS installation process. The BATWDA data set often is not created because it is so large, and the information is often too general for detailed performance analysis. CPExpert does not use the MICS BATWDA file. Rather, CPExpert uses data created by a proprietary modification to MICS. The CPExpert modification to MICS is described in this section.

The CPExpert modification to MICS allows you to collect essential information relating the DASD use to specific job steps and performance groups or service classes (with Goal Mode), without requiring massive amounts of DASD space. Additionally, the modification allows you to collect information required to relate I/O activity to specific data sets. The SAS data sets created by the modification to MICS normally are small (typically about 10 cylinders of DASD), and they are easy to process.

The CPExpert modification to MICS invokes two standard MICS user exits and includes a simple (one-line) modification to the DYSMFFMT module. Installing the modification for MICS is an optional step; however the DASD Component will be unable to relate DASD use and performance constraints to specific applications or to specific data sets unless you install the modification.

Implementing the CPExpert modification to MICS consists of five steps:

Step 1: Install the DASDMIC code.

Install the DASDMIC code into the prefix.MICS.USER.SOURCE library containing user modifications to MICS. This is done by copying the DASDMIC member from the CPEXPERT.SOURCE partitioned data set into the prefix.MICS.USER.SOURCE library. The DASDMIC member contains all the macros that will be invoked by the modifications described below.

Step 2: Modify the sharedprefix.MICS.USER.SOURCE(#SMFEXIT).

Modify the sharedprefix.MICS.USER.SOURCE(#SMFEXIT), by including SAS statements (1) to include the CPExpert macros, (2) to define the _USRDMAP macro, and (3) to define the _USRSSF macro.

Exhibit 2-5 illustrates the normal sharedprefix.MICS.USER.SOURCE(#SMFEXIT) member before the modifications. Exhibit 2-6 illustrates the normal sharedprefix.MICS.USER.SOURCE(#SMFEXIT) member after the modifications.

```
/* some comments */
%INCLUDE SOURCE(#SMFEXIT);
```

NORMAL sharedprefix.MICS.USER.SOURCE(#SMFEXIT) BEFORE MODIFICATION

EXHIBIT 2-5

```
/* some comments */
%INCLUDE SOURCE(#SMFEXIT);

/* CPEXPERT MODIFICATION */

%INCLUDE USOURCE(DASDMIC);

MACRO _USRDMAP
/* DEVICE ADDRESS MAPPING EXIT */
%%DASD_MIC;
%

 MACRO _USRSSFS
/*  STEP TERMINATION INTERIM SMF FILE EXIT */
%%CPEOUTDD;
%
```

NORMAL sharedprefix.MICS.USER.SOURCE(#SMFEXIT) AFTER MODIFICATION

EXHIBIT 2-6

WARNING: Please note that two (2) "%" characters are used when invoking "new style" macros inside "old style" macros in the user exits.

- The **%INCLUDE USOURCE(DASDMIC);** statement shown in Exhibit 2-6 causes MICS to include the DASDMIC macros as MICS loads its own code.

- The **%%DASD_MIC;** SAS macro statement shown as part of the _USRDMAP macro in Exhibit 2-6 generates code and arrays to process and store information related to job step use of DASD, by device address.

- The **%%CPEOUTDD;** SAS macro statement shown as part of the _USRSSFS macro in Exhibit 2-6 generates code to output the CPEDASD.DASD30DD data set, and clears the arrays described above.

The effect of Step 2 is that the normal _USRDMAP and _USRSSFS user exits included as null members in sharedprefix.MICS.SOURCE(#SMFEXIT) will be overridden by the code required for CPExpert.

Step 3: Modify prefix.MICS.SOURCE(DYSMFFMT).

Modify the sharedprefix.MICS.SOURCE(DYSMFFM1) module[1] by including the following SAS statement just after the "DATA" statement:

```
%VARDDCPE(MAXDASD=nn,MAXDD=nnn,MINIO=nnn,COLLECT=xxxxxxx)
```

WARNING: Do not include a semicolon after the macro invocation!

The %VARDDCPE macro statement generates code to define the DASD30DD data set, define the DASD30DS data set, and describe the variables contained in the KEEP statement for the data sets. The macro parameters specify the maximum number of unique DASD devices referenced by any job step in your environment, the maximum number of DD statements associated with critical data sets, the significant number of I/O operations directed to any data set in any Type 30(Interval) record, and which type of data collection should be performed.

- The **MAXDASD** macro parameter specifies the maximum number of unique DASD devices that are referenced by any job step in your environment. Note

[1] It is unfortunate that a modification must be made to the MICS coding, since this is contrary to standards implemented at many installations. A request was made to LEGENT Corporation (through the MICS Advisory Council, as Incident #441496) that a standard user exit be implemented after the DATA statement. If this request had been accepted and implemented, output files could be generated easily within the standard MICS user exit facility. Using a standard MICS user exit facility would obviate the need for a modification to MICS code. Computer Associates has not indicated a desire to make the modification.

that this is **not** the number of DD statements. Rather, this is the number of unique devices.

There often will be many DD statements for a single device, but the DASD_MIC software combines these DD statements into one observation in the DASD30DD data set.

The default specification for the MAXDASD macro parameter is **MAXDASD=25**. The DASD_MIC software will produce a warning message if this value is too low, and the software simply ignores devices for any job step exceeding the value specified. (There is no other effect; that is, the software will not abort and the analysis software will continue to function properly.)

The only reason for specifying 25 versus 200 (for example) as an initial value, is that the larger the value, the more DASD space that will be used for the DASD30DD data set. The record size will increase by 15 bytes for each unique device specified. If you process 20,000 job steps each day and specify a maximum of 25 unique devices per job step, the total temporary DASD space required for the DASD30DD data set will be about 25 cylinders (3380). This is not a large amount in any event, but there is no point in writing larger records than are needed. Therefore, MAXDASD=25 should be specified initially and this value should be increased only if required.

- The **MAXDD** macro parameter specifies the maximum number of unique DD names that may be associated with critical data sets. As described in Section 1 (Chapter 5.3), the CPExpert modification to MICS creates an array which contains the DD names of all DD statements associated with data sets which have been defined to CPExpert with a response objective. (The array is created only if you specify DDNAMES or BOTH in the COLLECT macro parameter described below.)

  Each array element requires only 8 bytes of storage. This amount storage is insignificant. However, SAS generates a SAS data element name for each array element. SAS 5.18 has a restriction on the total number of data element names which can be defined. When this number is exceeded, SAS aborts with an error in the compilation stage. Therefore, you cannot set an extremely large value for the MAXDD parameter when executing under SAS 5.18.

  The default specification for the MAXDD macro parameter is **MAXDD=200**, generating an array capable of containing 200 unique DD names. The DASD_MIC software will produce a warning message if the number of unique DD names is greater than the array size. The software will list all DD names which it could not process, and it simply ignores DD names for any job step exceeding the value specified. (There is no other effect; that is, the software will not abort and the analysis software will continue to function properly.) If you receive a warning message from the CPExpert modification, you should consider (1) increasing the value of the MAXDD macro parameter or (2) restricting the number of data sets you define as critical.

- The **MINIO** macro parameter specifies the minimum number of I/O operations to any data set within a particular SMF Type 30(Interval) record. The CPExpert modification will ignore any data sets encountered if the number of I/O operations is below the MINIO value. The point of this parameter is that you probably will not worry about data sets with a very low activity, regardless of the DASD response time they experienced. There is no point in collecting information about very low activity data sets.
  The default specification for the MINIO macro parameter is **MINIO=100**.

- The **COLLECT** macro parameter specifies whether to collect device information (required to associate device activity with application systems), to collect DD name information (required to associate device activity with data sets), or to collect both device and DD name information.

  Specify **COLLECT=DEVICES** to collect device information. This is the default setting. If you have previously installed the modification for MICS, you do not have to make any changes to your installation.

  Specify **COLLECT=DDNAMES** to collect DD name information.

  Specify **COLLECT=BOTH** to collect both device information and DD name information.

  WARNING: Please review Chapter 6 of this section before specifying DDNAMES or BOTH as an option for the COLLECT parameter. You should **not** specify either of these options unless you have defined critical data sets and have followed the procedures in Chapter 6.

Exhibit 2-7 illustrates the prefix.MICS.SOURCE(DYSMFFM1) code with the modification, reflecting explicit stating of the default values.

```
DATA
 %VARDDCPE(MAXDASD=25,MAXDD=200,MINIO=100,COLLECT=DEVICES)
 %MAINWRK(CCC=SMF,WRK=INI)
```

**SAMPLE sharedprefix.MICS.SOURCE(DYSMFFM1)
WITH CPEXPERT MODIFICATION**

EXHIBIT 2-7

Step 4: Add the CPEDASD DD statement to the JCL.

Add the CPEDASD DD statement to the job control language (JCL) used to execute MICS. The CPEDASD DD statement defines the SAS library in which the CPExpert modification will place the DASD30DD and DASD30DS data sets created by the modification.

Exhibit 2-8 illustrates a sample CPEDASD DD statement.

```
//CPEDASD DD DSN=prefix.CPEXPERT.CPEDATA,DISP=OLD
```

SAMPLE CPEDASD DD STATEMENT IN MICS JCL

EXHIBIT 2-8

The CPEDASD DD statement can refer to the CPEDATA SAS library if you wish, or you can allocate a different SAS library. For space allocation purposes the CPEDASD data set requires slightly less than one cylinder of DASD space per 2,200 SMF Type 30(Interval) records processed by MXG or MICS to contain the application-related information written to the DASD30DD SAS data set (that is, if you have specified COLLECT=DEVICES or COLLECT=BOTH). For example, if you typically have about 22,000 Type 30(Interval) records in your daily SMF file, you will need about 10 cylinders of DASD allocated to the CPEDASD SAS library to contain the application-related information.

It is difficult to estimate the amount of space required for the data set-related information written to the DASD30DS data set, since this is a function of the number of critical data sets you define and the number of unique DD names associated with the data sets. The DASD30DS data set contains about 730 observations per track.

MICS users may wish to place the CPEDASD DD statement into PROCLIB and then do a MICS JCLGEN to copy the information into the MICS prefix.CNTL libname.

We STRONGLY suggest that you test the modification in a MICS Test environment before placing the modification into a MICS production environment! If you (or we) have made a mistake, the MICS production run might abort. This certainly would not generate a favorable impression about CPExpert.

Chapter 3: Defining workload categories

The Workload Definition Section of the DASGUIDE module can be used to (1) define workload categories, (2) associate performance groups with the workload categories prior to MVS (Goal Mode), (3) associate service classes with the workload category with MVS (Goal Mode), (4) assign a relative importance to the workloads, and (5) direct CPExpert to perform analysis based upon the defined workload categories.

The workload definition is optional; you do not have define any workloads. However, you must define workloads if you wish CPExpert to analyze the DASD performance effects of contending workloads. Exhibit 2-9 illustrates a sample workload definition.

```
******************************************************************* ;
*  WORKLOAD DEFINITION BY PERFORMANCE GROUP                          ;
******************************************************************* ;
* DO NOT REMOVE THE FOLLOWING SAS MACRO COMMENT LINE!                ;
/* WORKLOAD DEFINITION
BEGIN THE WORKLOAD DEFINITION;
BATCH= 1             * SAMPLE: BATCH IS PGN 1                        ;
TSO  = 2,6,9         * SAMPLE: TSO IS PGN 2,6,9                      ;
CICS = 12,14         * SAMPLE:  CICS IS PGN 12,14                    ;
* DO NOT REMOVE THE FOLLOWING MACRO COMMENT LINE!                    ;
*/
******************************************************************* ;

******************************************************************* ;
*  WORKLOAD DEFINITION BY SERVICE CLASS
       ;
******************************************************************* ;
* DO NOT REMOVE THE FOLLOWING SAS MACRO COMMENT LINE!                ;
/* WORKLOAD DEFINITION (GOAL MODE) BEGIN THE WORKLOAD DEFINITION;
BATCH = BATHI,BATMED,BATLOW
CICS        = CICSAOR1,CICSAOR2,CICSAOR3
* DO NOT REMOVE THE FOLLOWING MACRO COMMENT LINE!                    ;
*/
```

SAMPLE DISPLAY OF CPEXPERT.USOURCE(DASGUIDE) MODULE

EXHIBIT 2-9

he workload category definition syntax **prior** to MVS (Goal Mode) is:

NAME = PGN1[,PGN2...,PGNn]

- **NAME** is the workload name (e.g., BATCH, TSO, CICS, etc.). The name an be any number of characters (including blanks), since it is delimited by the equal sign. However, you should generally restrict the length to 10 characters, since some reports produced by CPExpert assume a maximum length for the column containing the workload name. No more than 10 workload category names may be defined. A workload name must not begin with "*".

- **PGN** is the number of the performance group(s) associated with this workload category. The performance group numbers are separated by commas, and the entire set may optionally be enclosed in parentheses. Up to 10 performance groups may be associated with a workload category[2]. Any performance groups not associated with a workload category are placed in the OTHER category.

The workload category definition syntax **with** MVS (Goal Mode) is:

NAME = SRV1[,SRV2...,SRVn]

- **NAME** is the workload name (e.g., BATCH, TSO, CICS, etc.). The name an be any number of characters (including blanks), since it is delimited by the equal sign. However, you should generally restrict the length to 10 characters, since some reports produced by CPExpert assume a maximum length for the column containing the workload name. No more than 10 workload category names may be defined. A workload name must not begin with "*".

- **SRV** is the service class name(s) associated with this workload category. The service class names are separated by commas, and the entire set may optionally be enclosed in parentheses. Up to 10 service classes may be associated with a workload category[3]. Any service classes not associated with a workload category are placed in the OTHER category.

NOTE: If you are running Goal Mode, please be sure that you have modified USOURCE(GENGUIDE) to specify **%LET GOALMODE=Y;** so the CPExpert code related to service classes will be generated.

---

[2]The limits of 10 different workload categories and 10 performance groups per category were selected because they seem so large that they should satisfy your performance analysis needs. Please give us a call if you have a requirement for more than these limits.

[3]The limits of 10 different workload categories and 10 service classes per category were selected because they seem so large that they should satisfy your performance analysis needs. Please give us a call if you have a requirement for more than these limits.

Summary of Coding Restrictions:

- A workload name must **not** begin with "*".
- A maximum of 10 workload categories may be defined.
- A maximum of 10 performance groups or service classes may be assigned to any workload category.
- The entire definition for a workload category must be contained within a single line[4].
- You may define the same performance groups or service classes in more than one workload category. However, the analysis may give strange results if the workload categories are defined as "competing" with each other (you have the strange situation of a performance group or service class competing with itself!).

There is a potentially significant limitation with defining workload categories: workload categories can be identified only by performance group or service class rather than performance group or service class **period**. This limitation is imposed by the SMF data; the standard SMF data associates performance or resource use to performance group periods or service class periods only in the SMF Type 72 records. It is not possible to associate DASD use, for example, to specific performance group or service class periods.

[4]This did not seem to be a particularly restrictive design constraint, so it didn't seem necessary to incorporate "continuation line" logic. If this does pose a problem, please give Computer Management Sciences a call.

Chapter 4: Defining critical data sets

The Data Set Definition Section of the DASGUIDE module can be used to (1) define data sets and (2) associate a response objective with each data set.

The data set definition is optional; you do not have define any data sets. However, you must define data sets and specify a response objective for each data set if you wish CPExpert to analyze the DASD performance provided to critical data sets.

Exhibit 2-10 illustrates a sample data set definition.

```
*****************************************************************;
*  DATA SET NAME/RESPONSE OBJECTIVE DEFINITIONS
         ;
*****************************************************************;
* DO NOT REMOVE THE FOLLOWING SAS MACRO COMMENT LINE!            ;
/* DEFINE DATA SET NAMES                                ;
DSN=USER.EXTERNAL.DATA,RESPONSE=20                ;
DSN=USER.INTERNAL,RESPONSE=30   ;
DSN=USER.INTERNAL.EXECUTIVE.DATA,RESPONSE=5;
DSN=USER.INTERNAL.SYSTEMS.DATA,RESPONSE=20;
* DO NOT REMOVE THE FOLLOWING MACRO COMMENT LINE!          ;
*/
*****************************************************************;
```

SAMPLE DISPLAY OF CPEXPERT.USOURCE(DASGUIDE) MODULE

EXHIBIT 2-10

The data set definition syntax is:

DSN = data.set.name,RESPONSE=obj

- **DSN** is the data set name. The data set name can be any valid data set name. The processing of data set name "wild cards" and the order in which data set names are processed are described below.
- **RESPONSE** is the device response objective for the identified data set. This specification is in milliseconds. For example, RESPONSE=20 means that the

device on which the data set resides should provide an average I/O response of 20 milliseconds during any RMF measurement interval. A response objective must be specified for each data set defined.

CPExpert places each data set name into SAS macro variables, and places the associated response objective into macro variables corresponding to the data set name.

- CPExpert processes the data set names with an implicit trailing "wild card" specification. That is, if only high level qualifiers are specified, all data set names beginning with the high level qualifiers are considered to be associated with the specified response objective.

- The order of appearance in the list controls the application of response objectives.

Examine Exhibit 2-10 to appreciate the significance of the above rules.

- All data sets beginning with "USER.EXTERNAL.DATA" will have a response objective of 20 milliseconds.

- All data sets beginning with "USER.INTERNAL" will have a response objective of 30 milliseconds. However, any data set beginning with "USER.INTERNAL.EXECUTIVE.DATA" will have a response objective of 5 milliseconds, while any data set beginning with "USER.INTERNAL.SYSTEMS.DATA" will have a response objective of 20 milliseconds.

- If the statement "DSN=USER.INTERNAL,RESPONSE=30" in Exhibit 2-10 were to be placed at the end of the list, it would override the previous statements and all data sets beginning with "USER.INTERNAL" would have a response objective of 30 milliseconds, regardless of whether they were associated with "EXECUTIVE.DATA" or with "SYSTEMS.DATA".

WARNING: Do not specify response objectives for data sets unless they are critical to accomplishing your management objectives, and do not make excessive use of the "wild card" technique. The purpose of this warning is to alert you to the fact that significant processing time and DASD space could be required if you do not judiciously apply this analysis technique.

The following operational notes apply to defining critical data sets:

- The DAS1415 module of CPExpert must be executed before you execute MXG or MICS to perform your normal daily update of your performance data base. The DAS1415 module processes the data set name and response objectives you have defined in USOURCE(DASGUIDE). The DAS1415 module then processes the raw SMF Type 14/15 records to identify jobs which reference the data sets, and to build a SAS data set containing the DD names used to reference the data sets. Section 4 describes how to execute the DAS1415 module.

- You **cannot** add data set names to USOURCE(DASGUIDE) after you have executed MXG or MICS for any particular day, because these data set names will not have been processed by the DAS1415 module. Since the data set names were not processed by the DAS1415 module, the CPExpert modification to MXG or MICS would not select data set information while MXG or MICS updated your performance data base.

- You **can** delete data set names or change the response objectives for data set names after your performance data base has been updated. For example, you may wish to do this between successive runs of the DASD Component of CPExpert. The DASD Component will analyze DASD performance based upon current information in USOURCE(DASGUIDE). This is possible because the DASD Component processes USOURCE(DASGUIDE) to acquire current information. This information (data set names and response objectives) is used to guide the analysis of your performance data base.

- The VOLSER exclude or select logic described in Section 3 (Chapter 2.3 and Chapter 2.4, respectively) can be used to exclude or select specific volumes, without regard to the data set name processing.

Chapter 5: Defining Multiple PDBs

The Multiple Performance Data Base (PDB) section of the DASGUIDE module can be used to define multiple PDBs to CPExpert.

The multiple PDB definition is optional; you do not have use or define multiple PDBs. This feature is applicable only to users who wish to use the DASD Component to analyze shared DASD contention problems (see Section 2, Chapter 2.5) **and** you have multiple performance data bases that contain information relating to systems that share DASD.

Exhibit 2-11 illustrates a sample definition of multiple performance data bases.

```
*****************************************************************;
* MULTIPLE PERFORMANCE DATA BASE (PDB)  DEFINITIONS
          ;
*****************************************************************;
%LET MULTIPDB =Y;
%LET PDBLIB2 =    PDBLIBB;
%LET PDBLIB3 =    PDBLIBC;
*****************************************************************;
```

SAMPLE DISPLAY OF CPEXPERT.USOURCE(DASGUIDE) MODULE

EXHIBIT 2-11

You should specify **%LET MULTIPDB=Y;** in prefix.CPEXPERT.USOURCE(GENGUIDE) to tell CPExpert to process multiple performance data bases. You may then specify up to eight performance data bases in addition to the standard performance data base described by the PDBLIB DD statement, for a total of nine performance data bases.

You identify the DD statements that describe the additional performance data bases by specifying **%LET PDBLIBn = ddname;** , where “n” ranges from 2 through 9. In the example shown in Exhibit 2-11, the second performance data base is identified by the DD name of PDBLIBB, and the third performance data base is identified by the DD name of PDBLIBC. You can use any valid DD name to describe the performance data bases. The definitions must be in numerical order (that is, the first additional PDB must be described by PDBLIB2, the second additional PDB must be described by PDBLIB3, etc.).

Section 3: Specifying Guidance Variables

The CPEXPERT.USOURCE(DASGUIDE) partitioned data set member contains variables to establish the overall guidance for the DASD Component. You modify the variables in the DASGUIDE member whenever you wish to change the guidance to CPExpert. This chapter describes these variables, how the variables are used, and what should be specified for the variables.

The variables in the DASGUIDE module can be viewed as "data selection and presentation" variables and "analysis control" variables. These two types of control variables are discussed separately.

The data selection and presentation variables allow you to select particular time intervals to be analyzed, and allow you to specify how the results from the analysis are to be presented.

The analysis control variables allow you to control the analysis the DASD Component will perform. These variables specify whether expanded analysis is to be performed, identify critical workloads, specify critical data sets, exclude devices from analysis, control CPExpert's analysis of VSAM data set problems or potential problems, etc.

Please do not allow CPExpert to perform analysis or produce reports which are meaningless in your environment. If the analysis and reports produced by CPExpert do not meet your needs, alter the guidance to CPExpert. If the guidance is insufficient, please call Computer Management Sciences at (703) 922-7027 (or e-mail **Don_Deese@cpexpert.com**) so we can make changes to improve CPExpert for you!

Chapter 1: Data Selection and Presentation Variables

The data selection and presentation variables allow you to select particular time intervals to be analyzed, and allow you to specify how the results from the analysis are to be presented. This chapter describes these variables.

Exhibit 3-1 illustrates the data selection and presentation variables contained in USOURCE(DASGUIDE).

```
**********************************************************************;
*    DATA SELECTION AND PRESENTATION VARIABLES                        ;
**********************************************************************;
%LET DASDATES =01FEB1991;    * START DATE FOR DATA ANALYSIS            ;
%LET DASTIMES =08:00:00;     * START TIME FOR DATA ANALYSIS            ;
%LET DASDATEE =31DEC9999;    * END DATE FOR DATA ANALYSIS              ;
%LET DASTIMEE =16:00:00;     * END TIME FOR DATA ANALYSIS              ;
%LET DASDAT2E =0;            * DEFAULT SECOND SELECTION DATE - END     ;
%LET DASDAT2S =0;            * DEFAULT SECOND SELECTION DATE - START   ;
%LET DASTIM2E =0;            * DEFAULT SECOND SELECTION TIME - END     ;
%LET DASTIM2S =0;            * DEFAULT SECOND SELECTION TIME - START   ;
%LET SHIFT    =Y;            * START AND END TIMES REFER TO A SHIFT    ;
%LET SYSTEM   =*ALL;         * SPECIFY SYSTEM TO PROCESS (*ALL=ALL)    ;
%LET SYSTEMn  =system;       * PROCESS SYSTEMn (n = 1-9)               ;
%LET SYSPLEX  =*ALL;         * SPECIFY SYSPLEX TO PROCESS (*ALL=ALL)   ;
%LET CONFIG   =&CPEWORK;     * SAS LIBRARY FOR I/O CONFIGURATION DATA  ;
%LET SASODS   =N;            * CONTROLS WHETHER SAS ODS IS USED        ;
%LET PATH     =;             * PATH FOR ODS OUTPUT                     ;
%LET FRAME    =DASFRAME;     * GENERIC ODS FRAME NAME                  ;
%LET CONTENTS =DASDCONT;     * GENERIC ODS CONTENTS NAME               ;
%LET BODY     =DASDBODY;     * GENERIC ODS BODY NAME                   ;
%LET STYLE    =;             * ODS HTML STYLE OPTION                   ;
%LET PDFODS   =N;            * CONTROLS WHETHER SAS PDF IS USED        ;
%LET LINKPDF  =;             * LINK TO CPEXPERT DOCUMENTATION          ;
%LET URL      =N;            * CONTROLS .HTM IN SAS ODS FRAME OUTPUT   ;
%LET VERBOSE  =V;            * RESULTS: VERBOSE/CONCISE/SUMMARY        ;
%LET MAXRULES =1000;         * PRODUCE -REPORT TOO BIG- GUIDANCE       ;
```

DATA SELECTION AND PRESENTATION VARIABLES

EXHIBIT 3-1

Chapter 1.1: DASDATES and DASTIMES variables

The DASDATES and DASTIMES variables specify the start date and start time, respectively, for the interval of SMF data the DASD Component is to analyze. These variables (in conjunction with the DASDATEE and DASTIMEE variables) allow you to select specific periods of data to analyze. The following example shows how to specify that data selection should start at 08:00:00 on March 6, 2001:

```
%LET DASDATES = 06MAR2001;   * START DATE FOR DATA ANALYSIS;

%LET DASTIMES  = 08:00:00;       * START TIME FOR DATA ANALYSIS;
```

Chapter 1.2: DASDATEE and DASTIMEE variables

The DASDATEE and DASTIMEE variables specify the end date and end time, respectively, for the interval of SMF data the DASD Component is to analyze. The following example shows how to specify that data selection should end at 17:00:00 on March 8, 2001:

```
%LET DASDATEE = 08MAR2001  * END DATE FOR DATA ANALYSIS;

%LET DASTIMEE  = 17:00:00;       * END TIME FOR DATA ANALYSIS;
```

Chapter 1.3: DASDAT2S and DASTIM2S variables

The DASDAT2S and DASTIM2S variables are optional. These variables specify the start date and start time, respectively, for a second interval of SMF data the DASD Component is to analyze. These variables (in conjunction with the optional DASDAT2E and DASTIM2E variables) allow you to select a second period of data to analyze, in addition to the period specified by the DASDATES/DASTIMES and DASDATEE/DASTIMEE selection variables. The following example shows how to specify that a second period of data selection should start at 20:00:00 on March 6, 2001:

```
%LET DASDAT2S = 06MAR2001  * START DATE FOR DATA ANALYSIS;

%LET DASTIM2S  = 20:00:00;       * START TIME FOR DATA ANALYSIS;
```

Chapter 1.4: DASDAT2E and DASTIM2E variables

The DASDAT2E and DASTIM2E variables are optional. These variables specify the end date and end time, respectively, for a second interval of SMF data the DAS Component is to analyze. These variables (in conjunction with the optional DASDAT2S and DASTIM2S variables) allow you to select a second period of data to analyze, in addition to the period specified by the DASDATES/DASTIMES and DASDATEE/DASTIMEE selection variables. The following example shows how to specify that a second period of data selection should end at 22:00:00 on March 6, 2001:

```
%LET DASDAT2E = 06MAR2001;   * END DATE FOR DATA ANALYSIS;

%LET DASTIM2E  = 22:00:00;       * END TIME FOR DATA ANALYSIS;
```

Chapter 1.5: SHIFT variable

The SHIFT variable is used with the DASDATES, DASTIMES, DASDATEE, and DASTIMEE variables. The SHIFT variable allows you indicate how the time-selection variables should be used.

- If the SHIFT variable is "N", the time-selection will be based upon the absolute start and end dates/times specified. For example, if you wish CPExpert to process **all** data during a week, the start date and start time would be specified as the beginning of the week, and the end date and end time would be specified as the end of the week. You would specify "%LET SHIFT = N;" to process each 24-hour day. In the example shown above, data would be processed from 08:00:00 on 4 March until 17:00:00 on 8 March.

- If the SHIFT variable is "Y", the time-selection will be based upon the start and end dates, and the start and end times within each selected date. In the example shown above, perhaps you wished to process only the daily shift beginning at 08:00:00 and ending at 17:00:00. You would specify "%LET SHIFT = Y;" to process only the identified shift data, during the selected dates.

Chapter 1.6: SYSTEM variable

The SYSTEM variable is used to specify whether all systems in the performance data base should be evaluated, or to select a specific system identification to be evaluated.

Some users have data from multiple systems in their performance data base. For many of these users, or for users who have data for a single system represented in their performance data base, the default "*ALL" will be appropriate. No change of the SYSTEM variable would be required for these users.

However, some users who have data from multiple systems may wish to evaluate only a single system with the parameters specified in this member of DASGUIDE. For example, they might be temporarily interested in evaluating the performance of only an "important" system (such as a major production system) and not be interested in evaluating the performance of other systems with data in the performance data base. This evaluation can be accomplished by changing the SYSTEM variable to specify the system identification to be evaluated. For example, to specify that only data from SYS1 should be evaluated, specify:

```
%LET SYSTEM = SYS1 ; * PROCESS ONLY DATA FROM SYS1;
```

In another situation, a CPExpert user might wish to evaluate different systems with different DASGUIDE parameters. These different evaluations can be accomplished by different executions of the DASD Component. For each execution of the DASD Component, the USOURCE DD statement would be changed to reference different USOURCE libraries. Each USOURCE library would contain guidance members with appropriate guidance variables. The SYSTEM variable for each DASGUIDE guidance member would specify the system identification to which the guidance applied.

If you are using MICS as your performance data base, the SYSTEM variable refers to the MICS system identifier, rather than the SMF system identifier.

Chapter 1.7: SYSTEMn variable(s)

The SYSTEMn variable(s) are used to select multiple systems to be evaluated.

As described in the SYSTEM guidance variable discussion above, some sites have data from multiple systems in their performance data base. These sites can process data from all systems by specifying %LET SYSTEM=ALL; in USOURCE(CICGUIDE), or can select a specific system to process by specifying %LET SYSTEM=system; in USOURCE(CICGUIDE), where “system” is the system identification of the system to be processed.

Some sites have data from multiple systems in their performance data base and do not want to process all systems, but do wish to process more than one system. For example, some systems might be production systems and some might be test systems. For these sites, the **SYSTEMn** guidance variable can be used to select more than one specific system to analyze.

The SYSTEM guidance variable can be used to select data from only one system to analyze, and the SYSTEMn guidance variable(s) can be used to select up to 9 additional systems to analyze. For example, if you wish to analyze data from four systems (named SYSA, SYSB, SYSC, AND SYSX) in a single execution of the DASD Component, specify:

```
%LET SYSTEM  = SYSA ; * PROCESS DATA FROM SYSA;
%LET SYSTEM1 = SYSB ; * PROCESS DATA FROM SYSB;
%LET SYSTEM2 = SYSC ; * PROCESS DATA FROM SYSC;
%LET SYSTEM3 = SYSX ; * PROCESS DATA FROM SYSX;
```

Chapter 1.8: SYSPLEX variable

The SYSPLEX variable is used to specify whether data from each sysplex in the performance data base should be evaluated, or whether CPExpert should select data for a specific sysplex to be evaluated.

Some users have data from more than one sysplex in their performance data base. For many of these users, or for users who have data for a single sysplex represented in their performance data base, the default "*ALL" will be appropriate. No change of the SYSPLEX variable would be required for these users.

However, some users who have data from more than one sysplex may wish to evaluate only a single sysplex with the parameters specified in this member of DASGUIDE. This evaluation can be accomplished by changing the SYSPLEX variable to specify the sysplex to be evaluated. The following example shows how to specify that data only from PRODPLEX should be evaluated:

```
%LET SYSPLEX = PRODPLEX ; * PROCESS DATA ONLY FROM PRODPLEX;
```

Chapter 1.9: CONFIG variable

The CONFIG variable allows organizations to “save” the model of the I/O configuration that is created by CPExpert into a SAS library, described by the CONFIG DDNAME. The default value for the CONFIG guidance variable is **%LET CONFIG=WORK;** which causes the I/O configuration model to be placed into the SAS WORK temporary library. You can cause the DASD Component to place the I/O configuration into a permanent SAS library by using the CONFIG guidance variable. The I/O configuration model created by CPExpert contains not only the I/O configuration (physical channels, logical channels, control units, and devices attached to each system), but also contains key performance metrics related to each aspect of the I/O configuration. This option can be useful for reporting I/O activity for capacity planning. If you wish to save the I/O configuration model produced by CPExpert, specify **%LET CONFIG=ddname;**, where "ddname" is the name of JCL DD statement that describes the SAS library where the I/O configuration model is to be placed.

Chapter 1.10: SAS Output Delivery System

Output from CPExpert is created using Basic SAS statements. This Basic SAS output is designed for a standard SAS printer (line) format. With SAS Release 8, SAS users can use the SAS Output Delivery System to create output that is formatted in Hypertext Markup Language (HTML). This output can be browsed with Internet Explorer, Netscape, or any other browser that fully supports the HTML 3.2 tag set.

The CPExpert WLM Component, DB2 Component, CICS Component, and DASD Component support the SAS ODS features.

Please reference the *CPExpert Installation Guide* for more detailed information about using the SAS ODS feature of CPExpert.

Chapter 1.11: VERBOSE variable

The VERBOSE variable provides a control on the amount of narrative that the DASD Component lists with each rule result. Some installations prefer to produce only concise findings, and evaluate the results when the findings are significant. Other installations wish to produce expanded findings, and evaluate the results on a daily basis. You can use the VERBOSE variable to control the amount of narrative, depending upon your preferences. The options with the VERBOSE variable are:

V = Print verbose comments related to each rule that was produced during the evaluation of DASD performance. The verbose comments describe the rule, provide key information associated with the rule, and may provide a specific reference related to the rule.

C = Print "concise" results about the DASD devices. Each "rule record" created by the DASD Component has a "level" associated with it. "Level 1" records are those records describing basic information about the DASD devices with the most potential for performance improvement. "Level 2" records are those records which analyze the causes of poor performance and suggest ways in which improvements may be made. When the **C** verbose option is specified, only Level 1 information is printed.

S = Print "summary" results about the DASD devices. When the **S** verbose option is specified, only DAS000 (sysplex-wide intensity) and DAS050 (system intensity) information is printed.

Chapter 1.12: MAXRULES variable

A very large amount (several hundred pages) of output can be produced during some executions of the DASD Component. Such a large amount of output is not easy to analyze, and tends to cause the report to be rejected by analysts or management. Fortunately, there are several options in the DASD Component that can be used to reduce the amount of output. New users of the DASD Component might not be aware of these options, and might not be favorably impressed with the DASD Component since they would perceive the DASD Component output to be useless. This is not a desirable result!

The MAXRULES variable is used to control whether CPExpert detects that "a large amount of output" was produced. CPExpert produces suggestions on how to reduce the output to a more manageable size when the number of rules (and supporting data) exceed the value of the MAXRULES guidance variable.

The default value of the MAXRULES guidance variable is 1000, which should be approximately 50 pages of output (basic rules have an average of 10 lines per description, while supporting rules have only one line for the data). You can alter the MAXRULES guidance variable and cause CPExpert to produce the suggestions for ways to reduce the output when either more or less rules (and supporting data) are produced. For example, to indicate that the suggestions on ways to reduce the output should be printed only when more than 2000 rules were produced, specify:

```
%LET MAXRULES=2000 ; * PRODUCE "REPORT TOO BIG" GUIDANCE;
```

Chapter 2: Analysis Control Variables

The analysis control variables allow you to control how the DASD Component analyzes your DASD configuration. These variables can be used to (1) specify whether expanded analysis is to be performed based upon workload categories, (2) specify whether expanded analysis is to be performed based upon data set names, (3) exclude volumes from analysis altogether, or (4) select specific volumes for analysis.

The analysis control variables are optional. You do not have to specify any analysis control variables unless you wish to alter the basic processing performed by the DASD This chapter describes the analysis control variables associated with the DASD Component. Exhibit 3-2 illustrates the analysis control variables contained in USOURCE(DASGUIDE). Please note that these are basic analysis guidance variables. Additional guidance variables are listed in other chapters in this section.

```
*********************************************************************;
*   ANALYSIS GUIDANCE VARIABLES                                     ;
*********************************************************************;
%LET ANALYZE   = 3;    * DEFAULT FOR NUMBER OF DEVICES TO ANALYZE      ;
%LET DASDEXCP  =100;   * MINIMUM NUMBER OF EXCP TO SELECT (DASD30DD)   ;
%LET DASDSN    = N;    * ANALYZE CRITICAL DATA SETS?                   ;
%LET EVALDASD  =Y;     * OPTION TO PRODUCE ONLY I/O CONFIGURATION      ;
%LET EXCLUDE   = N;    * DEFAULT FOR EXCLUDING VOLUMES                 ;
%LET GOALMODE  = Y;    * OPERATING IN GOAL MODE                        ;
%LET LIST42DS  =10;    * LIST ONLY TOP NUMBER OF DATA SETS             ;
%LET LISTALL   =Y;     * LIST DATA FOR ALL RMF INTERVALS;
%LET LOVED1    =N;     * PERFORM 'LOVED ONE' ANALYSIS?                 ;
%LET MIN42PCT  =0.1;   * EXCLUDE REPORTING LOW-ACTIVITY DATA SETS      ;
%LET MINIORT   =0      * MINIMUM I/O RATE TO ANALYZE                   ;
%LET MINIOWT   =0;     * MINIMUM I/O INTENSITY (OR WEIGHT) TO ANALYZE  ;
%LET MINRESP   =0;     * MINIMUM I/O RESPONSE TO ANALYZE               ;
%LET REPORT    =11;    * NUMBER OF TOTAL VOLUMES TO REPORT             ;
%LET SELECT    = Y;    * DEFAULT FOR VOLUME SELECTION                  ;
%LET SHARED    = N;    * DEFAULT FOR PROCESSING SHARED SYSTEM DATA     ;
%LET TYPE30DD  = N;    * MXG TYPE30 MODIFICATION (Y = AVAILABLE)       ;
%LET TYPE42DS  = N;    * MXG TYPE42DS (Y = AVAILABLE)                  ;
*********************************************************************;
```

ANALYSIS GUIDANCE VARIABLES

EXHIBIT 3-2

Chapter 2.1: Number of devices to analysis: ANALYZE variable

The optional ANALYZE variable is used to specify the number of devices which CPExpert should analyze in detail.

As explained in Section 1, most Data Base Administrators are able to address a limited number of DASD problems each day. The DASD Component was designed to identify the most serious DASD problems. The initial design identified the device with the most potential for improvement, and analyzed this device in detail.

Some users of CPExpert felt that the "identify the device with the most potential for improvement" design philosophy was unnecessarily restrictive. These users wanted to specify the number of devices which should be analyzed in detail. The ANALYZE variable meets this requirement.

You can direct CPExpert to analyze up to 999 DASD devices in a single execution, by using the ANALYZE variable to specify the number of devices to analyze in detail. CPExpert will provide a detailed analysis of all devices whose device response time exceeded the average device response time for their device type, up to the number specified by the ANALYZE variable. For example, if you wish CPExpert to analyze the top 25 devices with the worst performance, specify **%LET ANALYZE=25;** in USOURCE(DASGUIDE).

You likely will receive an unwieldy amount of output if you specify a large number of devices using the ANALYZE variable. One way to limit the output is to specify **%LET VERBOSE=S** so that only device summary results will be printed.

Chapter 2.2: Exclude devices with low activity: DASDEXCP variable

The modification to MXG or NeuMICS includes a MINIO parameter which specifies the minimum number of I/O operations to any data set within a particular SMF Type 30(Interval) record. The CPExpert modification will ignore any data sets encountered if the number of I/O operations is below the MINIO value. The default for the MINIO parameter is MINIO=100, indicating that data sets will be included in the DASD30DD data set generated by the CPExpert modification if at least 100 EXCPs were directed to the data set.

This screening is based on data set, and is not based on the total I/O activity to the device. Consequently, some devices might have data represented in the DASD30DD data set (generated by the CPExpert modification), but these devices might have very low activity. For example, a device could be represented in the DASD30DD with only 100 EXCPs per SMF Type 30 recording interval if only a single data set were referenced 100 times.

It normally is not interesting to consider devices with a very low activity, and many sites have a large number of such devices in the DASD30DD data set created by CPExpert. To reduce processing time, the DASDEXCP guidance variable can be used to exclude such low-activity volumes.

The default value for the DASDEXCP guidance variable is 100 (the same as the MINIO variable in the CPExpert modification) that creates the DASD30DD data set. If you wish CPExpert to ignore devices with low activity, change to DASDEXCP guidance to a larger value. For example, if you wish CPExpert to consider only those devices with more than 1000 EXCPs per SMF Type 30 recording interval, specify **%LET DASDEXCP=1000;** in USOURCE(DASGUIDE).

Chapter 2.3: Analyze using response objectives: DASDSN variable

The DASGUIDE **DASDSN** guidance variable specifies whether to perform expanded analysis based upon data set response objectives. Specify **%LET DASDSN = Y;** to tell CPExpert to perform expanded analysis based upon data set response objectives.

Please refer to Section 2 (Chapter 4) for operational notes regarding changing your specification of critical data sets or changing response objectives associated with the data sets.

Please refer to Section 3 (Chapter 5) for a description of the procedures to follow if you wish to perform analysis based on data set response objectives.

Chapter 2.4: Produce only I/O configuration: EVALDASD variable

The EVALDASD guidance variable allows organizations to suppress the evaluation of DASD I/O problems. This option is for those users who wish to use the DASD Component simply to produce SAS data sets representing the I/O configuration (and associated data) for each system in the sysplex, and a SAS data set for the overall sysplex.

The DASD Component processes SMF data (in a performance data base) to create a single record for each RMF measurement interval, that contains information about the channel paths, controllers, and device for each volume. This information is interesting to some users for management reporting purposes, or for data base administrator reporting purposes. These users do not necessarily want to use the DASD Component to analyze DASD problems, since their sole use of the DASD Component is to have an easy-to-use tool that creates the I/O configuration information. This option can be useful for reporting I/O activity for capacity planning.

- If the EVALDASD guidance variable is "Y" (which is the default value), the DASD Component will create the model of the I/O configuration and evaluate DASD performance problems represented by this model.

- If the EVALDASD guidance variable is "N", the DASD Component will create the model of the I/O configuration and place the results in the SAS library described by the optional **CONFIG DDNAME** DD statement. The DASD Component will then issue an **ABORT RETURN 100**; statement within the SAS coding. This statement results in SAS aborting the SAS session, with a return code of 100.

Chapter 2.5: Excluding volumes from analysis: EXCLUDE variable

The EXCLUDE guidance variable is used to tell CPExpert whether you wish to exclude volumes from analysis. The EXCLUDE variable acts as a "switch" to control whether CPExpert processes the DASGUIDE member searching for volumes to exclude from analysis. The point of having a "switch" variable is that some installations may wish to regularly exclude volumes from analysis, but periodically analyze all volumes.

If the EXCLUDE variable is **N**, CPExpert will not exclude any volumes from analysis. If the EXCLUDE variable is **Y**, CPExpert will process the DASGUIDE member to identify all volumes to be excluded.

Please refer to Chapter 3 for more information on excluding volumes from analysis.

Chapter 2.6: Specifying data sets to list: LIST42DS variable

If SMF Type 42 (Data Set Statistics) information is available[1] in a MXG performance data base, the DASD Component will process MXG TYPE42DS to select data sets that were referenced during RMF measurement intervals in which the poorly performing devices exceeded the average performance. There can be many data sets referenced on the devices (hundreds or even thousands of data sets can be referenced). It is not helpful to have a large number of data sets listed. Consequently, CPExpert uses the **LIST42DS** variable to limit the number of data sets listed in any RMF measurement interval, for a particular device. Only the number of data sets specified by the LIST42DS variable will be listed individually, and any remaining data sets will be summarized and listed on a single line.

The default value for the LIST42DS variable is 10, indicating that 10 data sets will be listed for each RMF interval in which a poorly-performing device had a performance problem.

You can alter this number of data sets listed by using the LIST42DS guidance variable. For example, if you wish CPExpert to list only 5 data sets for each RMF interval in which a poorly-performing device had a performance problem, specify **%LET LIST42DS=5;** in USOURCE(DASGUIDE).

Chapter 2.7: Perform “loved one” analysis: LOVED1 variable

The LOVED1 guidance variable is used to control whether the DASD Component performs “loved one” analysis.

[1] **%LET TYPE42DS = Y;** must be specified in USOURCE(GENGUIDE) or USOURCE(DASGUIDE) to advise CPExpert that TYPE42DS is available.

IBM speakers at professional conferences have long advocated the performance analysis principle of "Know your loved ones and always have someone to kick around." As described in Section 1 (Chapter 5.2) and Section 2 (Chapter 5), you can identify workload categories (such as TSO, CICS, BATCH, etc.) to the DASD Component. The DASD Component can then analyze your DASD configuration from the view of the workload categories, focusing on the performance provided to your "loved one" workload.

Analysis of critical (or "loved one") workload is possible only if you accomplish the following:

- You have installed the modification to MXG or MICS to collect application-related information as described in Section 2 of this User Manual. The DASD Component cannot perform expanded analysis based on workload categories unless it has information relating DASD use to specific performance groups or service classes, jobs, and job steps. This information is extracted from the SMF Type 30(DD) records by the modifications described in Chapter 1 and Chapter 2 of Section 2.

- You have defined workload categories to the DASD Component. At least one workload category (the most critical category) must be defined using the process described in Chapter 3 of Section 2.

- You identify the critical (or "loved one") workload to the DASD Component using the **LOVED1** guidance variable.

- You specify **%LET TYPE30DD = Y;** in the USOURCE(GENGUIDE) CPExpert general guidance module. This tells CPExpert that you have installed the modification to MXG or MICS to collect application-related information.

If you accomplish the above steps, the DASD Component will perform an expanded analysis of your DASD configuration from the perspective of the "loved one" workload category. That is, only devices and channel paths used by the critical workload will be analyzed. This analysis can be quite useful to determine which DASD devices are providing the worst performance from the perspective of the critical workload.

The **LOVED1** guidance variable specifies whether to perform expanded analysis of workload categories, and specifies the most critical (or "loved one") workload.

For example, if your critical workload is CICS production executing in Service Class CICSAOR1, Service Class CICSAOR2, and Service Class CICSAOR3 you might define the CICS workload as illustrated in Exhibit 2-9 in Section 2. You could identify the CICS workload as your most critical workload as:

```
%LET LOVED1 = CICS ; * CICS IS THE CRITICAL WORKLOAD;
```

In this example, the DASD Component will analyze all DASD use by Service Class CICSAOR1, CICSAOR2, and CICSAOR3. The DASD Component will isolate the devices providing the worst performance from the view of this CICS workload, identify likely causes of poor performance, identify the job and job steps **in other service classes** which **probably** cause performance to be poor, and suggest ways to correct the problems. A similar analysis would be performed for performance groups, if you are operating in Compatibility Mode.

A significant phrase was used in the above paragraph: "probably cause performance to be poor". Please review the discussion associated with appropriate rules in Appendix A to appreciate why the identification might not be precise, but might be only "probable". In brief summary, the identification is "probable" rather than "precise" because of synchronization problems between SMF Type 30 data and SMF Type 70(series) data, unless you have specified that the SMF/RMF recording be synchronized. This issue is discussed in some detail in Section 5 and in appropriate rules in Appendix A.

The DASD Component will not perform "loved one" analysis if the LOVED1 macro variable is null (this is the default specification).

The "loved one" analysis is incompatible with the data set response objective analysis[2]. That is, if you perform "loved one" analysis, you cannot perform an analysis of response objectives for critical data sets in the same execution of CPExpert. You could perform "loved one" analysis in one execution of the DASD Component and perform data set response objective analysis in another execution of the DASD Component, processing the same data in your performance data base.

Chapter 2.8: Exclude reporting low-activity data sets: MIN42PCT variable

If SMF Type 42 (Data Set Statistics) information is available[3] in a MXG performance data base, the DASD Component will process MXG TYPE42DS to select data sets that were referenced during RMF measurement intervals in which the poorly performing devices exceeded the average performance. There can be many data sets referenced on the devices (hundreds or even thousands of data sets can be referenced). It is not helpful to have a large number of data sets listed. Additionally, some data sets might have a low I/O activity and would not be interesting to analyze. Consequently, CPExpert uses the **MIN42PCT** variable to limit the number of data sets listed in any RMF measurement interval, for a particular device. Only data sets having a percent of activity for the total

[2]This design decision was made because there would be significant problems with interpreting the results from CPExpert if both types of analysis were performed simultaneously, and because a large coding effort would be required to perform the concurrent analysis. Please call Computer Management Sciences if you have a strong requirement for concurrent analysis.

[3]%LET TYPE42DS = Y; was specified in USOURCE(GENGUIDE).

volume greater than the percent specified by the MIN42PCT variable will be listed individually, and any remaining data sets will be summarized and listed on a single line.

The default value for the MIN42PCT variable is 0.1, indicating that data sets will not be listed individually for any RMF interval unless the data set intensity of access (I/O rate * response time) was greater than 0.1% of the total volume intensity of access. You can alter this percent of data sets listed by using the MIN42PCT guidance variable.

For example, if you specified **%LET MIN42PCT = 25;** in USOURCE(DASGUIDE), a maximum of 4 data sets would be listed in any RMF interval since no more than 4 data sets could have 25% or higher access intensity.

Please note that, regardless of the data set access intensity of any particular data set, only the number of data sets specified by the **LIST42DS** guidance variable (described earlier) will be listed. This means that there are two ways to limit the number of data sets listed: (1) the LIST42DS which limits the number of data sets listed in any RMF interval, and (2) the **MIN42PCT** guidance variable which limits the data sets listed to those that exceed the specified percent.

Chapter 2.9: List data for all RMF intervals: LISTALL variable

Once the DASD Component has selected one or more devices to analyze, it analyzes device-related data for all RMF intervals. This analysis determines whether each device has performance problems, and determines the most significant problems during the RMF interval. When the DASD Component produces the results of its analysis, it shows device information for each RMF interval analyzed, regardless of whether a performance problem existed with the device. This listing of all RMF intervals is the default, since it presents a continuous view of the device performance for all RMF intervals in the performance data base.

Some users analyze DASD problems for many systems in a single execution of the DASD Component, and they collect RMF data on short time intervals (for example, 15 minute RMF intervals or shorter). The default listing of data for all RMF intervals can produce very large reports for these users.

The LISTALL variable can be used to suppress the listing of RMF intervals in which no performance problem was detected with a device. You can suppress listing all RMF intervals by specifying **%LET LISTALL=N;** in USOURCE(DASGUIDE). With this specification, the DASD Component will list only those intervals in which performance problems were detected.

Chapter 2.10: Minimum I/O rate to analyze: MINIORT variable

Before the DASD Component selects a device for analysis, the code determines whether the device has an I/O rate greater than the MINIORT value. This feature allows user to exclude low-activity devices from analysis. The default value for the MINIORT variable is 0, indicating that all devices will be candidates for analysis, regardless of their I/O rate. You can alter this I/O rate by using the MINIORT guidance variable.

For example, if you specified **%LET MINIORT = 1;** in USOURCE(DASGUIDE), only devices with a minimum I/O rate greater than 1 I/O per second would be selected for analysis.

Chapter 2.11: Minimum I/O rate to analyze: MINIOWT variable

Before the DASD Component selects a device for analysis, the code determines whether the device has an I/O "intensity" (or weight) greater than the MINIOWT value. This feature allows user to exclude low-activity devices from analysis. The default value for the MINIOWT variable is 0, indicating that all devices will be candidates for analysis, regardless of their I/O intensity. You can alter this I/O rate by using the MINIOWT guidance variable.

For example, if you specified **%LET MINIORT = 1000;** in USOURCE(DASGUIDE), only devices with a minimum I/O intensity greater than 1000 would be selected for analysis.

Chapter 2.12: Minimum I/O response to analyze: MINRESP variable

Before the DASD Component selects a device for analysis, the code determines whether the device has an I/O response greater than the MINRESP value. This feature allows user to exclude low-activity devices from analysis. The default value for the MINRESP variable is 0, indicating that all devices will be candidates for analysis, regardless of their I/O response. You can alter this I/O rate by using the MINRESP guidance variable. The MINRESP variable is specified in millisecond units.

For example, if you specified **%LET MINRESP = 5;** in USOURCE(DASGUIDE), only devices with a minimum I/O response greater than milliseconds would be selected for analysis.

Chapter 2.13: Number of volumes to report: REPORT variable

The DASD Component analyzes performance information about the device(s) with the most performance improvement potential, based on ANALYZE guidance variable described earlier. To put the performance improvement potential of these devices in perspective relative to the overall sysplex or system, the DASD Component also produces a brief summary of relevant characteristics of other devices, with the list ranked in descending order by device access intensity (I/O RATE * average I/O response). The summary by sysplex is produced in Rule DAS000 and the summary by system is produced in Rule DAS050.

The number of devices for which summary information is produced is controlled by the **REPORT** guidance variable. The default for the REPORT guidance variable is 10, indicating that information is produced for the top 10 devices on a system basis. The value of the REPORT variable is doubled for the sysplex view.

Chapter 2.14: Selecting volumes to analyze: SELECT variable

The SELECT guidance variable is used to tell CPExpert whether you wish to select specific volumes for analysis. The SELECT variable acts as a "switch" to control whether CPExpert processes the DASGUIDE member searching for volumes to select for analysis. The point of having a "switch" variable is that some installations may wish to regularly analyze all volumes, but periodically select specific volumes for analysis.

If the SELECT variable is **N**, CPExpert will not select specific volumes for analysis, but will analyze all volumes after applying data selection criteria. If the SELECT variable is **Y**, CPExpert will process the DASGUIDE member to identify the volumes to be analyzed.

Please refer to Chapter 4 for more information on selecting volumes to analyze.

Chapter 2.15: Analyze shared DASD Conflicts - SHARED variable

CPExpert can perform an analysis of conflicts between DASD shared between systems or MVS images. The shared DASD analysis is accomplished when you perform the following steps:

- You collect and process SMF data for each system to be analyzed. The SMF data must be placed into your standard performance data base. This step is not specific for CPExpert; it is the normal processing you would do. The step is listed first just for completeness.

- **You should ensure that the system clocks on all systems are set at roughly the same time**. CPExpert will examine the SMF data contained in your performance data base, based upon the SMF time associated with the measurement data. Shared DASD problems between systems will be detected based on this SMF time. If the SMF times are significantly different, the analysis might not properly identify conflicts or conflicting applications.

You should not worry that the system clocks be set at **exactly** the same time. The analysis performed by CPExpert is not intended to identify an isolated performance problem. Rather, CPExpert attempts to identify those problems which **continually** cause shared DASD performance problems. If a shared DASD problem is continual, the problem will be reflected in multiple SMF recording intervals. Consequently, it is not essential that the SMF times be exactly matched between systems.

If the shared devices are in the same sysplex, this step will be performed automatically by the sysplex timer. The step is listed just for completeness.

- You specify **%LET SHARED = Y;** in USOURCE(DASGUIDE) to tell CPExpert to perform analysis of shared DASD.

- If you wish the DASD Component to identify applications causing shared DASD conflicts, you must perform the following additional steps:

  - Install the modification to MXG or MICS to collect data set-related information as described in Section 2 of this User Manual.

  - Specify **%LET TYPE30DD = Y;** in the USOURCE(GENGUIDE) general guidance module. This tells CPExpert that you have installed the modification to MXG or NeuMICS to collect application-related information.

Chapter 2.16: SMF Type 30 modification installed - TYPE30DD variable

The **TYPE30DD** guidance variable tells CPExpert whether you have installed the CPExpert modification to MXG or MICS code to collect SMF Type 30 (Data Definition) information[4].

This modification allows CPExpert to relate DASD performance information contained in SMF Type 74 records to specific service classes or performance groups, and to relate the Type 74 information to specific jobs or job steps.

The TYPE30DD statement should be specified as **%LET TYPE30DD = Y;** if you have installed the CPExpert modification to collect Type 30 (Data Definition) information for the system(s) being analyzed.

[4]The TYPE30DD guidance variable is normally specified when installing CPExpert and is discussed in the *CPExpert Installation Guide*. The TYPE30DD guidance variable is included in this document simply to remind you that the SMF Type 30 information is required if you wish to have CPExpert show applications that access devices with poor performance, wish to perform "loved one" analysis, or wish to perform critical data set analysis. The TYPE30DD is not included in the USOURCE(DASGUIDE) member released with CPExpert, since a specification in USOURCE(DASGUIDE) would override any specification in USOURCE(GENGUIDE).

Chapter 2.17: SMF Type 42 (Data Set Statistics)- TYPE42DS variable

The **TYPE42DS** guidance variable tells CPExpert whether SMF Type 42 (Data Set) records are available for analysis[5]. If you collect the SMF Type 42 (Data Set) records, CPExpert can report data set information[6] for data sets residing on volumes with poor performance.

The default value for the TYPE42DS guidance variable is "N", indicating that the TYPE42DS data set is not available for analysis. You can change this guidance if Type 42 (Data Set) records are available, and if you wish CPExpert to report data set information, by specifying **%LET TYPE42DS=Y;** in USOURCE(DASGUIDE).

[5] The TYPE42DS data set statistics information is available only with a MXG performance data base.

[6] The TYPE42DS guidance variable is normally specified when installing CPExpert and is discussed in the *CPExpert Installation Guide*. The TYPE42DS guidance variable is included in this document simply to remind you that the SMF Type 42 information is required if you wish to have CPExpert show data sets that are accessed on devices with poor performance. The TYPE42DS is not included in the USOURCE(DASGUIDE) member released with CPExpert, since a specification in USOURCE(DASGUIDE) would override any specification in USOURCE(GENGUIDE).

Chapter 3: Excluding volumes from analysis

For a variety of reasons, you may wish to exclude certain volumes from analysis. The most common reason is that CPExpert repeatedly identifies problems with particular volumes, but (1) you do not wish to make changes to correct the problems, (2) you are unable to make changes (because of application requirements or political realities), or (3) the DASD Component analysis is "flawed" because of data problems or data averaging. Whatever the reason, you may wish to exclude certain volumes from analysis.

Excluding volumes from analysis is optional; you do not have to exclude any volumes from analysis. If you do **not** exclude volumes from analysis, CPExpert will analyze every DASD VOLSER encountered in SMF Type 74 data after applying data selection criteria.

If you wish to exclude volumes from analysis, you must (1) use the **EXCLUDE** guidance variable and (2) define the volumes to exclude.

Chapter 3.1: EXCLUDE variable

The EXCLUDE guidance variable is used to tell CPExpert whether you wish to exclude volumes from analysis. The EXCLUDE variable acts as a "switch" to control whether CPExpert processes the DASGUIDE member searching for volumes to exclude from analysis. The point of having a "switch" variable is that some installations may wish to regularly exclude volumes from analysis, but periodically analyze all volumes.

If the EXCLUDE variable is **N**, CPExpert will not exclude any volumes from analysis. If the EXCLUDE variable is **Y**, CPExpert will process the DASGUIDE member to identify all volumes to be excluded.

Chapter 3.2: Defining volumes to exclude

You may exclude any number of volumes from analysis by coding the name of the volume in the "EXCLUDE VOLUME" portion of the DASGUIDE Module. You simply enter the VOLSER of each volume to be excluded after the "EXCLUDE" statement.

- The volumes to be excluded are entered one VOLSER per line or multiple VOLSERs per line.

- If you include multiple VOLSERs on a single line, you must separate them by blanks or commas.

- You can include comments on the line by placing "*" before the comment.

- You can use "generic" exclude logic by simply listing the first "n" characters of the VOLSERs you wish to exclude. For example, to exclude all VOLSERs beginning with ABC, simply specify ABC as the VOLSER to exclude.

CPExpert will process the DASGUIDE Module, searching for the "EXCLUDE" statement. Any VOLSERs between the "EXCLUDE" statement and the "*/" statement will be placed into global SAS macro variables and may be excluded from analysis.

Exhibit 3-3 was extracted from the distributed USOURCE(DASGUIDE) module, and illustrates how to exclude volumes from analysis.

```
***************************************************************;
* OPTIONAL GUIDANCE TO EXCLUDE VOLUMES FROM ANALYSIS;
***************************************************************;
* DO NOT REMOVE OR ALTER THE FOLLOWING SAS MACRO COMMENT LINE!;
 /* EXCLUDE THE FOLLOWING VOLSER FROM ANALYSIS:
EXCLUDE
 MVS21AA   * SAMPLE VOLSER DEFINITION
 CHKPT01   * SAMPLE VOLSER DEFINITION
 CHPPT02   * SAMPLE VOLSER DEFINITION
 SPOOL1,SPOOL2,SPOOL3
* DO NOT REMOVE THE FOLLOWING MACRO COMMENT LINE!
 */
```

EXCLUDING SPECIFIC VOLUMES FROM ANALYSIS

EXHIBIT 3-3

Chapter 4: Selecting specific volumes for analysis

You may wish to select specific volumes for analysis by the DASD Component, irrespective of whether these volumes are among the "worst" performing devices. For example, you may wish to examine only volumes containing specific application data sets.

Selecting volumes for analysis is optional; you do not have to select specific volumes for analysis. If you do not select specific volumes for analysis, CPExpert will analyze every DASD VOLSER encountered in SMF Type 74 data after applying data selection criteria. If you **do** select specific volumes for analysis, CPExpert will report on only those volumes (but CPExpert will analyze the performance of the volumes within the context of the overall I/O configuration).

If you wish to select volumes from analysis, you must (1) use the **SELECT** guidance variable and (2) define the volumes to select.

Chapter 4.1: SELECT variable

The SELECT guidance variable is used to tell CPExpert whether you wish to select specific volumes for analysis. The SELECT variable acts as a "switch" to control whether CPExpert processes the DASGUIDE member searching for volumes to select for analysis. The point of having a "switch" variable is that some installations may wish to regularly analyze all volumes, but periodically select specific volumes for analysis.

If the SELECT variable is **N**, CPExpert will not select specific volumes for analysis, but will analyze all volumes after applying data selection criteria. If the SELECT variable is **Y**, CPExpert will process the DASGUIDE member to identify the volumes to be analyzed.

Chapter 4.2: Defining volumes to analyze

You may select any number of volumes for specific analysis by coding the name of the volume in the "SELECT VOLSER" portion of the DASGUIDE Module. You simply enter the VOLSER of each volume to be excluded after the "SELECT" statement.

- The volumes to be selected are entered one VOLSER per line or multiple VOLSERs per line.
- If you include multiple VOLSERs on a single line, you must separate them by blanks or commas.
- You can include comments on the line by placing "*" before the comment.

- You can use "generic" select logic by simply listing the first "n" characters of the VOLSERs you wish to select. For example, to select all VOLSERs beginning with DEF, simply specify DEF as the VOLSER to select.

CPExpert will process the DASGUIDE Module, searching for the "SELECT" statement. Any VOLSERs between the "SELECT" statement and the "*/" statement will be placed into global SAS macro variables and will be the only volumes CPExpert analyzes in detail (note that CPExpert will continue to process data describing the entire I/O configuration so that the analysis is done within the context of your configuration).

Exhibit 3-4 was extracted from the distributed USOURCE(DASGUIDE) module, and illustrates how to select specific volumes for analysis.

You should note that CPExpert still applies the "worst volume" philosophy when it reports the results from the analysis. That is, only volumes exceeding the average of those selected will be analyzed in detail. You might be confused when you examine the "next worst volumes" report, unless you keep this in mind. (Give us feedback if you don't like this approach. If it is unsatisfactory, we can change the design.)

```
*****************************************************************;
* OPTIONAL GUIDANCE TO SELECT SPECIFIC VOLUMES FOR ANALYSIS;
*****************************************************************;
* DO NOT REMOVE OR ALTER THE FOLLOWING SAS MACRO COMMENT LINE!;
 /* SELECT THE FOLLOWING VOLSER FOR ANALYSIS:
SELECT
 PAGE21    * SAMPLE VOLSER DEFINITION
 PAGE22    * SAMPLE VOLSER DEFINITION
 PAGE23    * SAMPLE VOLSER DEFINITION
 PAGE24    * SAMPLE VOLSER DEFINITION
* DO NOT REMOVE THE FOLLOWING MACRO COMMENT LINE!
 */
```

SELECTING SPECIFIC VOLUMES FOR ANALYSIS

EXHIBIT 3-4

Chapter 5: Analyzing response objectives for critical data sets

As described in Section 1 (Chapter 5.3: Expanded analysis - Specific Data Sets) and Section 2 (Chapter 4: Defining Critical Data Sets), you can identify critical data sets to the DASD Component and specify a response objective for these data sets. The DASD Component can then analyze the performance of your DASD configuration, in one of two ways: (1) analyzing data set response based on information from TYPE42DS statistics or (2) analyzing data based on information in TYPE14/15 and using the CPExpert modification to MXG or NeuMICS as the SMF Type30 records are processed.

Chapter 5.1: Analysis based on TYPE42DS

This method can provide comprehensive information, without requiring the modification to MXG. This method **is not applicable** to NeuMICS, since NeuMICS does not retain sufficient information related to SMF Type 42 (Data Set Statistics).

MXG creates TYPE42DS from the SMF Type 42 (Subtype 6 - data set I/O statistics) records created by SMS. The TYPE42DS file contains I/O access characteristics information, at the data set level. CPExpert extracts data set information and data set response statistics from TYPE42DS, and compares these to the response objectives you have made in USOURCE(DASGUIDE) about critical data sets. When any data set response time exceeds the specified objective, the DASD Component selects the data set and the volume it resides on for detailed analysis.

The following steps are necessary to implement the critical data set analysis based on TYPE42DS records:

- Specify **%LET DASDSN = Y;** in USOURCE(DASGUIDE) to tell the DASD Component to analyze DASD performance based upon data set name.

- Specify **%LET TYPE42DS = Y;** in the USOURCE(GENGUIDE) CPExpert general guidance module. This tells CPExpert that you have SMF Type 42 (Data Set Statistics) available in your performance data base.

- Define critical data sets to the DASD Component and specifie a response objective for each data set. This process is described in Section 2 (Chapter 6).

Chapter 5.2: Analysis based on TYPE14/15 and CPExpert modification

This method is a bit more involved, but can be used regardless of whether data sets are managed by SMS, and this method can be used regardless of whether you use MXG or NeuMICS to create your performance data base.

With this method, the DASD Component analyzes SMF Type 14/15 records to extract data set names that correspond to the critical data sets that you have identified in USOURCE(DASGUIDE). The CPExpert modification to MXG or NeuMICS extracts DD information as SMF Type 30 records are processed. The DASD Component then correlates the data set information with the DD information, to determine whether critical data sets exceed the specified response objective. When any data set responses time exceeds the specified objective, the DASD Component selects the data set and the volume it resides on for detailed analysis.

The following steps are necessary to implement the critical data set analysis based on SMF Type 14/15, Type 30, and Type 74

- Install the modification to MXG or NeuMICS to collect data set-related information as described in Section 2 of this User Manual. The DASD Component cannot perform expanded analysis based on specific data sets unless it has information relating DASD use to specific data sets, service classes, jobs, and job steps. This information is extracted from the SMF Type 30(DD) records by the modifications described in Section 2 (Chapter 3 and Chapter 4).

- Define critical data sets to the DASD Component and specified a response objective for each data set. This process is described in Section 2 (Chapter 6).

- Specify **%LET DASDSN = Y;** in USOURCE(DASGUIDE) to tell the DASD Component to analyze DASD performance based upon data set name.

- Specify **%LET TYPE30DD = Y;** in the USOURCE(GENGUIDE) CPExpert general guidance module. This tells CPExpert that you have installed the modification to MXG or MICS to collect application-related information.

- Collect SMF Type 14/15 records and execute the DAS1415 module (described in Section 4 of this manual) to process the SMF Type 14/15 records. You must execute the DAS1415 module **before** you execute your normal daily update of your performance data base.

Chapter 6: Analyzing VSAM data sets

VSAM data set activity typically accounts for a large percent of I/O activity (more than 70% at some sites). Tuning of a few files or correcting common problems often can result in significantly improved performance (IBM benchmarks show up to 90% improvement resulting from some simple changes).

Analysis of VSAM data sets is not applicable to performance data bases created with NeuMICS, since the required SMF Type 42(Data Set statistics) and Type 64 (VSAM statistics) are not available.

With CPExpert Release 12.2, the DASD Component was enhanced to provide a rudimentary analysis of common VSAM problems. Additional analysis will be added in future enhancements to the DASD Component.

The DASD Component can optionally analyze VSAM data set performance problems or potential problems, only if the MXG TYPE64 file is available (and CPExpert is provided with the **%LET TYPE64=Y;** guidance variable). Additionally, most analysis depends on having the MXG TYPE42DS file available (and CPExpert is provided with the **%LET TYPE42DS=Y;** guidance variable).

Most sites have many VSAM data sets. Some of the VSAM data sets are open for a long time, but most are open for only a short interval. Some of the VSAM data sets have a significant amount of I/O activity, while other VSAM data sets have very low activity. These characteristics can result in very large reports produced by the DASD Component, since the DASD component can report on problems with each VSAM data set.

Options are provided with the VSAM analysis to summarize findings for VSAM data sets, to ignore VSAM data sets that are open for a short time, or to ignore VSAM data sets that have less than a specified amount of I/O activity. Additionally, each finding is based on guidance variables. These guidance variables can be altered so CPExpert will produce results only when more restrictive conditions are met. All these options can be (and should be) used to reduce the output from the DASD Component's to a moderate size.

This chapter describes the guidance variables that are available for analyzing performance problems of VSAM data sets. Exhibit 3-5 illustrates the VSAM analysis guidance variables contained in USOURCE(DASGUIDE) Component.


```
**********************************************************************;
*   VSAM ANALYSIS GUIDANCE VARIABLES                                  ;
**********************************************************************;
%LET ANALVSAM  =BAD;   * VSAM ANALYSIS OPTION                           ;
%LET CASPLITS  = 10;   * EXCESSIVE CONTROL AREA SPLITS                  ;
%LET DIRINDEX  =25;    * PCT DIRECT ACCESS TO VSAM INDEX COMPONENT      ;
%LET EXTENTS   = 1;    * EXCESSIVE SECONDARY EXTENTS                    ;
%LET LSRSEQ    = 50;   * LSR SEQUENTIAL ACCESS DOMINATE                 ;
%LET MXEXTENT  =225;   * MAXIMUM EXTENTS INDICATING POTENTIAL PROBLEM   ;
%LET NSRDIR    = 50;   * NSR DIRECT ACCESS DOMINATE                     ;
%LET OPENTIME  =900;   * MINIMUM VSAM DATA SET OPEN TIME FOR ANALYSIS   ;
%LET PCTDIR    =80;    * PERCENT DIRECT VSAM ACCESSES                   ;
%LET PCTSEQ    =80;    * PERCENT SEQUENTIAL VSAM ACCESSES               ;
%LET VSAMEXCL  = NO;   * OPTION: EXCLUDE SELECTED VSAM DATA SETS        ;
%LET VSAMIO    = 100;  * VSAM I/O ACTIVITY SIGNIFICANT TO ANALYZE       ;
%LET VSAMSMRY  = NO;   * VSAM DATA SET SUMMARIZED?                      ;
***********************************************************************;
```

VSAM ANALYSIS GUIDANCE VARIABLES

EXHIBIT 3-5

Chapter 6.1: Controlling analysis of VSAM: ANALVSAM variable

The DASD Component optionally analyzes VSAM data sets to identify performance problems or potential problems if SMF Type 64 information is available, and if **%LET TYPE64=Y;** has been specified in USOURCE(DASGUIDE)[7].

For flexibility, there are four options that control whether and how the VSAM data set analysis is performed. These options are specified by the ANALVSAM guidance variable:

- **%LET ANALVSAM = BAD; is specified in USOURCE(DASGUIDE)**. This is the default specification. When **%LET ANALVSAM=BAD;** is specified, the analysis is done **only** for VSAM data sets that reside on the devices that have been selected for analysis because the devices are the "worst performing" DASD devices in the configuration.

[7] Most of the VSAM analysis also requires TYPE42DS information. Analysis that requires TYPE42DS information will be suppressed unless **%LET TYPE42DS=Y;** has been specified in either USOURCE(GENGUIDE) or USOURCE(DASGUIDE).

- **%LET ANALVSAM = ALL; is specified in USOURCE(DASGUIDE)**. This option directs CPExpert to analyze all VSAM data sets in the configuration, regardless of which devices the VSAM data sets reside. CPExpert will analyze VSAM performance for all volumes, after analyzing those volumes whose performance is worse than average. The basic data selection criteria (for example, date and time, system, etc.) will be applied before performing the VSAM analysis.

- **%LET ANALVSAM = ONLY; is specified in USOURCE(DASGUIDE)**. This option directs CPExpert to analyze **only** VSAM data sets, based on SMF Type 64 and (optionally) SMF Type 42 (Data Set Statistics) records. When the ONLY option is selected, CPExpert will analyze VSAM performance without performing the basic DASD analysis of worst performing volumes. Exercising this option eliminates the processing associated creating a model of the I/O configuration and analyzing device performance, and CPExpert will ignore devices with non-VSAM data sets. This option will be particularly useful when new applications are produced and these applications use VSAM data sets.

- **%LET ANALVSAM = NO; is specified in USOURCE(DASGUIDE)**. This option directs CPExpert to eliminate the VSAM analysis. This option might be selected after CPExpert has analyzed VSAM data set performance, and you have made all of the changes that you intend to implement.

Chapter 6.2: Excessive Control Area splits: CASPLITS variable

CPExpert examines the SMF Type 64 information contained in the MXG TYPE64 data set to identify VSAM data sets that have excessive Control Area splits. CPExpert sums the ACCASPLT variable (the number of CA splits since the data set was created) and the CASPLITS variable (the number of CA splits with the current OPEN). CPExpert compares this sum with the **CASPLITS** guidance variable in USOURCE(DASGUIDE). CPExpert produces Rule DAS600 when the total number of CA splits exceeds the value specified by the **CASPLITS** guidance variable.

The default value for the **CASPLITS** guidance variable is 10, indicating that CPExpert should produce Rule DAS600 when a VSAM data set experienced more than 10 CA splits. You should alter this guidance variable if you feel that Rule DAS600 is produced too often, or if you do not wish to take action when only 10 CA splits occur. For example, if you wish CPExpert to produce Rule DAS600 only when more than 50 CA splits occur for any VSAM data set, specify **%LET CASPLITS=50;** in USOURCE(DASGUIDE).

Chapter 6.3: Percent direct access to VSAM index component: DIRINDEX variable

CPExpert examines the SMF Type 64 information contained in MXG TYPE64 data set to identify VSAM KSDS or VRRDS data sets that have insufficient buffers assigned to the index component. CPExpert uses the TYPE42DS information to compute the percent of direct accesses to the VSAM data set, using the following algorithm:

$$\textit{Percent direct accesses} = \frac{S42AMDRB}{S42AMSRB + S42AMDRB}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces *Rule DAS621* under the following conditions:

- The TYPE42DS S42DSBUF variable showed that NSR was used for KSDS or VRRDS VSAM data sets, **and**

- The percent of direct accesses for the index component was greater than **DIRINDEX** guidance variable in USOURCE(DASGUIDE), **and**

- The number of buffers (the MXG BUFDRNO variable) assigned to the index component was less than the number of index levels (the MXG ACCLEVEL variable) for the VSAM data set.

CPExpert produces *Rule DAS622* under the following conditions:

- The TYPE42DS S42DSBUF variable showed that NSR was used for KSDS or VRRDS VSAM data sets, **and**

- The percent of direct accesses for the index component was greater than **DIRINDEX** guidance variable in USOURCE(DASGUIDE), **and**

- The STRNO specification in the ACB (the MXG ACBSTRNO) was greater than one (indicating that concurrent accesses had been specified for direct processing), **and**

- The number of buffers (the MXG BUFDRNO variable) assigned to the index component was less than the ACBSTRNO value, plus 1.

The default value for the DIRINDEX guidance variable is 25%, so CPExpert will produce Rule DAS621 or DAS622 for NSR VSAM data sets when more than 25% of the accesses were direct for the index component and the number of buffers assigned to the index component was less than the number of index levels (for rule DAS621) or less than the

STRNO value in the ACB, plus 1 (for Rule DAS622). You can change the DIRINDEX guidance variable if you feel that Rule DAS621 or Rule DAS622 are produced too often, or if you do not wish to take action when to increase the number of buffers assigned to the VSAM data sets listed.

For example, if you wish CPExpert to produce Rule DAS621 or Rule DAS622 only when more than 50% of the accesses for an index component were direct, specify **%LET DIRINDEX=50;** in USOURCE(DASGUIDE). You can completely suppress Rule DAS621 and Rule 622 by specifying **%LET DIRINDEX=100;** in USOURCE(DASGUIDE), since the percent cannot exceed 100.

Chapter 6.4: Excessive EXTENTS were allocated: EXTENTS variable

CPExpert examines the SMF Type 64 information contained in MXG TYPE64 data set to identify VSAM KSDS or VRRDS data sets that have excessive secondary allocations.

CPExpert compares NREXTNTS variable (the number of secondary extents in the VSAM data set this OPEN) with the **EXTENTS** guidance variable in USOURCE(DASGUIDE). CPExpert produces Rule DAS604 when the NREXTENT (the total number of extents) is greater than one, and the number of secondary extents allocated for this OPEN exceeds the value specified by the EXTENTS guidance variable.

The default value for the EXTENTS guidance variable is 0, indicating that CPExpert should produce Rule DAS604, Rule DAS605, or Rule DAS606 (depending on data encountered) when any secondary extent was allocated for the VSAM data sets listed. You can change the EXTENTS guidance variable if you feel that Rule DAS604 is produced too often, or if you do not wish to take action when secondary extents are allocated. For example, if you wish CPExpert to produce Rule DAS604 only when more than 5 secondary extents were allocated for VSAM data sets, specify **%LET EXTENTS=5;** in USOURCE(DASGUIDE). You can completely suppress Rule DAS635 by specifying **%LET EXTENTS=255;** in USOURCE(DASGUIDE), since VSAM cannot allocate 255 extents.

Chapter 6.5: Specifying LSR sequential access domination: LSRSEQ variable

With Local Shared Resource (LSR), VSAM buffers normally are shared among VSAM data sets accessed by tasks in the same address space. Since LSR is oriented toward shared (and direct) access, there is an expectation that a record might be re-used. Consequently, buffer management algorithms retain buffers as long as possible, using a least-recently used (LRU) algorithm, after a record is processed from the LSR buffers. LSR is not suited for applications that use sequential or skip sequential as their primary access mode, because there is no read-ahead algorithm with LSR, and there is no inherent overlap of

I/O and CPU processing. Consequently, using LSR for sequential access processing could degrade rather than improve performance.

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information to compute the percent of sequential accesses to the VSAM data set, using the following algorithm:

$$\text{Percent sequential accesses} = \frac{S42AMSRB}{S42AMSRB + S42AMDRB}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces Rule DAS635 under the following conditions:

- The TYPE42DS S42DSBUF variable showed that LSR was used for KSDS or VRRDS VSAM data sets, and

- The percent of sequential accesses for the data component was greater than the value specified for the **LSRSEQ** guidance variable in USOURCE(DASGUIDE.

The default value for the LSRSEQ guidance variable is 75%, so CPExpert will produce Rule DAS635 when LSR was specified as the buffering technique, and more than 75% of the accesses were sequential for the data component.

You can change the percent of sequential access that CPExpert uses to determine whether to produce Rule DAS635 by altering the LSRSEQ guidance variable. For example, if you wish CPExpert to produce Rule DAS635 only when more than 90% of the accesses to a LSR buffer pool were sequential, specify **%LET LSRSEQ=90;** in USOURCE(DASGUIDE). You can completely suppress Rule DAS635 by specifying **%LET LSRSEQ=100;** in USOURCE(DASGUIDE), since the percent cannot exceed 100.

Chapter 6.6: Specifying maximum extents: MXEXTENT variable

CPExpert examines the SMF Type 64 information contained in the MXG TYPE64 data set to identify VSAM data sets are in danger of reaching the maximum allowed number of extents. CPExpert compares NREXTENT variable (the total number of extents in the VSAM data set) with the **MXEXTENT** guidance variable in USOURCE(DASGUIDE). CPExpert produces Rule DAS607 when the NREXTENT is greater than the value specified by the MXEXTENT guidance variable **and** at least one extent was allocated during the current OPEN of the data set (CPExpert uses the NREXTNTS variable in TYPE64 for this decision).

The default value of the MXEXTENT guidance variable is 225, indicating that CPExpert should produce Rule DAS607 when at least 225 extents have been allocated for a VSAM data set. Since the maximum allowable is 255, the default value provides a threshold at which CPExpert provides notification that there is a potential problem.

You can alter this threshold by using the MXEXTENT guidance variable. For example, if you wish to be notified when the number of extents reach 200 (and at least one extent was required with the current OPEN), specify **%LET MXEXTENT=200;** in USOURCE(DASGUIDE).

Chapter 6.7: Specifying NSR direct access domination: NSRDIR variable

Non-shared resource (NSR) is the default VSAM buffering technique. VSAM data sets with NSR buffering can be accessed sequentially or direct (or both). However, NSR is suited for sequential processing because, if the data set access is sequential, the buffers are managed with a read-ahead algorithm. The read-ahead algorithm provides overlap of I/O and CPU processing and is efficient for sequential accesses. Since NSR is oriented toward sequential access, there is no expectation that a record will be re-used (as might exist with direct processing). Consequently, once a record is processed from the NSR buffers, the buffer is likely to be reclaimed for another record read from DASD.

NSR is not suited for direct processing, although NSR often **is** used for direct processing because it is easy to use and is the default buffering technique. Nonetheless, performance can be significantly improved if LSR is used for direct processing of VSAM data sets.

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information to compute the percent of direct accesses to the VSAM data, using the following algorithm:

$$\textit{Percent direct accesses} = \frac{\textit{S42AMDRB}}{\textit{S42AMSRB} + \textit{S42AMDRB}}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces Rule DAS625 under the following conditions:

- The TYPE42DS S42DSBUF variable showed that NSR was used for KSDS or VRRDS VSAM data sets, and

- The percent of direct accesses for the data component was greater than the value specified for the NSRDIR guidance variable in USOURCE(DASGUIDE.

The default value for the NSRDIR guidance variable is 75%, so CPExpert will produce Rule DAS625 when NSR was specified as the buffering technique, and more than 75% of the accesses were direct for the data component.

You can change the percent of sequential access that CPExpert uses to determine whether to produce Rule DAS635 by altering the NSRDIR guidance variable. For example, if you wish CPExpert to produce Rule DAS625 only when more than 90% of the accesses to a VSAM data set with NSR were sequential, specify **%LET NSRDIR=90;** in USOURCE(DASGUIDE). You can completely suppress Rule DAS625 by specifying **%LET NSRDIR=100;** in USOURCE(DASGUIDE), since the percent cannot exceed 100.

Chapter 6.8: Minimum VSAM open time: OPENTIME variable

Many installations have a large number of VSAM data sets that are open for a relatively short time, and analyzing these data sets would unnecessarily clutter the reports produced by the DASD Component. The **OPENTIME** guidance variable allows installations to control which data sets are analyzed, based on the amount of elapsed time that the data sets are open.

The SMF variables reflecting when the data set was opened (SMF30TM and SMF30DT) were available with z/OS Version 1 Release 1 (z/OS V1R1). For versions of MVS prior to z/OS V1R1, the default value for the OPENTIME guidance variable is zero, and this value cannot be changed (since CPExpert would have nothing in SMF to compare against). Beginning with z/OS V1R1, the default value for the OPENTIME guidance variable is 300, indicating that CPExpert should ignore[8] VSAM data sets that are open for less than 300 seconds (or 5 minutes).

If you are running MVS with z/OS V1R2 or subsequent release, and you wish CPExpert to ignore VSAM data sets that are open for different elapsed time, you can change the OPENTIME guidance variable to a different number of seconds. For example, if you wish CPExpert to analyze only those VSAM data sets that are open for more than 30 minutes, specify **%LET OPENTIME=1800;** in USOURCE(DASGUIDE).

[8]CPExpert does NOT ignore any VSAM data set based on open time if **%LET VSAMSMRY=Y;** has been specified in USOURCE(DASGUIDE), until all activity for each VSAM data set has been summarized. If **%LET VSAMSMRY=Y;** has been specified, the guidance is applied to the summarized data.

Chapter 6.9: Specifying percent direct access for control interval size: PCTDIR variable

A *Control Interval* is a continuous area of direct access storage that VSAM uses to store logical records. The size of Control Intervals can vary from one VSAM data set to another, but all the Control Intervals within the data portion of a particular data set must be the same length. The type of processing that is used should guide the choice of control interval size. When direct processing accounts for most of the accesses, a small data Control Interval is preferable. This is because only one record is retrieved at a time with direct processing. If a large Control Interval is specified, unnecessary I/O overhead is incurred reading the excess information. IBM suggests that a 4096 byte Control Interval normally would be appropriate for direct processing.

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information to compute the percent of accesses to the data component of the VSAM data set that were direct, using the following algorithm:

$$\mathit{Percent\ sequential\ accesses} = \frac{\mathit{S42AMSRB}}{\mathit{S42AMSRB} + \mathit{S42AMDRB}}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces Rule DAS611 when the percent of direct accesses for the data component was greater than the **PCTDIR** guidance variable in USOURCE(DASGUIDE), and the Control Interval size was greater than 4096 bytes. Additionally, CPExpert verifies that the maximum logical records (maximum LRECL) is less than 50% of the Control Interval size. This verification is done to make sure that Rule DAS611 is not produced for VSAM data sets that have spanned records.

The default value for the PCTDIR guidance variable is 80%, so CPExpert will produce Rule DAS611 when more than 80% of the accesses were direct for the data component, and the Control Interval size was more than 4096 bytes. You can alter this algorithm by changing the PCTDIR guidance variable in USOURCE(DASGUIDE). For example, if you wish CPExpert to produce Rule DAS611 only when more than 90% of the accesses to the data component were direct, specify **%LET PCTDIR=90;** in USOURCE(DASGUIDE). You can completely suppress Rule DAS611 by specifying **%LET PCTDIR=100;** in USOURCE(DASGUIDE), since the percent cannot exceed 100.

Chapter 6.10: Specifying percent sequential access for control interval size: PCTSEQ variable

A *Control Interval* is a continuous area of direct access storage that VSAM uses to store logical records. The size of Control Intervals can vary from one VSAM data set to another, but all the Control Intervals within the data portion of a particular data set must be the same length. The type of processing that is used should guide the choice of control interval size. When sequential processing accounts for most of the accesses, a large data Control Interval would normally be a good choice. This is because multiple records can be read into buffers and processed sequentially. For example, given a 16KB data buffer space, it is better to read two 8 KB Control Intervals with one I/O operation than four 4 KB Control Intervals with two I/O operations.

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information to compute the percent of sequential accesses to the data component of the VSAM data set, using the following algorithm:

$$Percent\ sequential\ accesses = \frac{S42AMSRB}{S42AMSRB + S42AMDRB}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces Rule DAS610 when the percent of sequential accesses for the data component was greater than the **PCTSEQ** guidance variable in USOURCE(DASGUIDE), and the Control Interval size was less than 8192 bytes.

The default value for the PCTSEQ guidance variable is 80%, so CPExpert will produce Rule DAS610 when more than 80% of the accesses were sequential for the data component, and the Control Interval size was less than 8192 bytes. You can alter this algorithm by changing the PCTSEQ guidance variable in USOURCE(DASGUIDE). For example, if you wish CPExpert to produce Rule DAS610 only when more than 90% of the accesses to the data component were sequential, specify **%LET PCTSEQ=90;** in USOURCE(DASGUIDE). You can completely suppress Rule DAS610 by specifying **%LET PCTSEQ=100;** in USOURCE(DASGUIDE), since the percent cannot exceed 100.

Chapter 6.11: Excluding VSAM data sets: VSAMEXCL variable

The **VSAMEXCL** guidance variable tells CPExpert whether you wish to exclude certain VSAM data sets from analysis. The VSAMEXCL variable acts as a "switch" to control whether CPExpert processes the DASGUIDE member searching for VSAM data set names to exclude from analysis. The point of having a "switch" variable is that some installations may wish to regularly exclude VSAM data sets from analysis, but periodically analyze all VSAM data sets.

If the VSAMEXCL variable is **N**, CPExpert will not exclude any VSAM data sets from analysis[9]. If the VSAMEXCL variable is **Y**, CPExpert will process the DASGUIDE member to identify all VSAM data sets to be excluded.

Please refer to Chapter 7 for more information on excluding VSAM data sets from analysis.

Chapter 6.12: Specifying significant VSAM I/O activity: VSAMIO variable

Many installations have a large number of VSAM data sets that have little I/O activity, and analyzing these data sets would unnecessarily clutter the reports produced by the DASD Component. The **VSAMIO** guidance variable allows installations to control which data sets are analyzed, based on the amount of I/O activity of the data sets. CPExpert examines the EXCP variable (number of EXCPs) in MXG TYPE64 records. CPExpert ignores[10] all records that have less than the amount of I/O activity specified in the VSAMIO guidance variable.

The default value for the VSAMIO guidance variable is 100, indicating that CPExpert should ignore VSAM data sets that have less than 100 I/O operations. This low value was selected as a default so that you could appreciate potential problems with most VSAM data sets in your installation.

If you wish CPExpert to ignore VSAM data sets that have a different I/O activity, you can change the VSAMIO guidance variable. For example, if you wish CPExpert to analyze only those VSAM data sets that have more than 1000 EXCPs during each OPEN, specify **%LET VSAMIO=1000;** in USOURCE(DASGUIDE).

[9] Please note that the ANALVSAM guidance variable also acts to exclude VSAM data sets from analysis. If **%LET ANALVSAM=BAD**; is specified in USOURCE(DASGUIDE), only VSAM data sets residing on poorly performing volumes will be analyzed.

[10] CPExpert does NOT ignore any VSAM data set based on I/O activity if **%LET VSAMSMRY=Y;** has been specified in USOURCE(DASGUIDE), until all activity for each VSAM data set has been summarized. If **%LET VSAMSMRY=Y;** has been specified, the guidance is applied to the summarized data.

Chapter 6.13: Summarizing VSAM activity: VSAMSMRY variable

Many installations have a large number of VSAM data sets that are open for a very short time, or have VSAM data sets that exhibit little I/O activity. Analyzing these data sets would unnecessarily clutter the reports produced by the DASD Component and produce a large amount of output[11]. The **VSAMSMRY** guidance variable can be used to cause CPExpert to summarize VSAM findings.

If **%LET VSAMSMRY=N;** is specified in USOURCE(DASGUIDE), CPExpert will list every occasion in which a finding applies to a specific VSAM data set. This listing will show the time that the VSAM information was written to SMF, based on merging of MXG TYPE64 and TYPE42DS records. The information will show the JOB NAME using the VSAM data set, and other information associated with the rule. Since many VSAM data sets are OPENed and CLOSEd frequently (particularly with CLOSE=T), a very large number of lines of output can be produced. However, this level of output might be useful if VSAM problems are to be correlated with specific times of poor performance.

If **%LET VSAMSMRY=Y;** is specified in USOURCE(DASGUIDE), CPExpert will summarize information for each VSAM data set and apply the VSAM analysis algorithms to the summarized information. The listing will show the time that the **last** VSAM information was written to SMF, based on merging of MXG TYPE64 and TYPE42DS records. The information will show the **last** JOB NAME using the VSAM data set, and a summary of other information associated with the rule. This level of output is useful for analyzing whether problems occur with VSAM data sets, over the entire measurement interval retained in the performance data base.

[11]Hundreds of pages of output might be created by CPExpert if every VSAM data set were analyzed each time that it was OPENed.

Chapter 7: Excluding VSAM data sets from analysis

For a variety of reasons, you may wish to exclude certain VSAM data sets from analysis. The most common reason is that CPExpert repeatedly identifies problems with particular VSAM data sets, but (1) you do not wish to make changes to correct the problems, (2) you are unable to make changes (because of application requirements or political realities), or (3) the DASD Component analysis is "flawed" because of data problems or data averaging. Whatever the reason, you may wish to exclude certain VSAM data sets from analysis.

Excluding VSAM data sets from analysis is optional; you do not have to exclude any VSAM data sets from analysis. If you do **not** exclude VSAM data sets from analysis, CPExpert will analyze every VSAM data set encountered in SMF Type 64 after applying data selection criteria.

If you wish to exclude VSAM data sets from analysis, you must (1) use the **VSAMEXCL** guidance variable and (2) define the VSAM data sets to exclude.

Chapter 7.1: VSAMEXCL variable

The VSAMEXCL guidance variable is used to tell CPExpert whether you wish to exclude VSAM data sets from analysis. The VSAMEXCL variable acts as a "switch" to control whether CPExpert processes the DASGUIDE member searching for VSAM data sets to exclude from analysis. The point of having a "switch" variable is that some installations may wish to regularly exclude VSAM data sets from analysis, but periodically analyze all VSAM data sets.

If the VSAMEXCL variable is **N**, CPExpert will not exclude any VSAM data sets from analysis. If the VSAMEXCL variable is **Y**, CPExpert will process the DASGUIDE member to identify all VSAM data sets to be excluded.

Chapter 7.2: Defining VSAM data sets to exclude

You may exclude any number of VSAM data sets from analysis by coding the name of the VSAM data set in the "EXCLUDE SPECIFIC VSAM DATA SETS FROM ANALYSIS" portion of the DASGUIDE Module. You simply enter the VSAM data set name of each VSAM data set to be excluded after the "/* EXCLUDE VSAM DATA SETS FROM ANALYSIS" statement.

- The VSAM data sets to be excluded are entered one data set name per line or multiple data set names per line.

- If you include multiple data set names on a single line, you must separate them by blanks or commas.

- You can include comments on the line by placing " *" before the comment. NOTE that a blank character MUST precede the asterisk.

- You can use "generic" exclude logic by simply listing the first "n" characters of the VSAM data set names you wish to exclude, followed by an asterisk. For example, to exclude all VSAM data set names beginning with “D10.RMF.MONITOR3" simply specify “D10.RMF.MONITOR3*" as the data set name to exclude.

CPExpert will process the DASGUIDE Module, searching for the "/* EXCLUDE VSAM DATA SETS FROM ANALYSIS" statement. Any VSAM data sets between the "/* EXCLUDE VSAM DATA SETS FROM ANALYSIS" statement and the "*/" statement will be placed into global SAS macro variables and may be excluded from analysis.

Exhibit 3-6 was extracted from the distributed USOURCE(DASGUIDE) module, and illustrates how to exclude VSAM data sets from analysis.

```
***********************************************************************;
* OPTIONAL GUIDANCE TO EXCLUDE SPECIFIC VSAM DATA SETS FROM ANALYSIS   ;
***********************************************************************;
* DO NOT REMOVE OR ALTER THE FOLLOWING SAS MACRO COMMENT LINE!         ;
/* EXCLUDE VSAM DATA SETS FROM ANALYSIS
DSN=RLSADSW.VF05D.ITEMACT.DATA
DSN=RLSADSW.VF07D.ITEMACT
DSN=D10.RMF.MONITOR3*
* DO NOT REMOVE THE FOLLOWING MACRO COMMENT LINE!
 */
***********************************************************************;
```

EXCLUDING VSAM DATA SETS FROM ANALYSIS

EXHIBIT 3-6

Section 4: Executing the DASD Component

This section describes how to execute the DASD Component of CPExpert.

The instructions in this section assume that you have installed the CPExpert software. The DASD Component is installed as normal part of installing CPExpert, and the modification to MXG or to MICS is installed as described in Section 2 of this User Manual. If you have not installed CPExpert and the modification to MXG or NeuMICS, please install the software before continuing.

Executing the DASD Component involves executing the DASCPE Module. Additionally, the DAS1415 Module must be executed if you wish to analyze the DASD performance provided to critical data sets.

Chapter 1: Executing the DASCPE Module

This chapter describes how to execute the DASCPE Module of the DASD Component.

As stated in the Introduction to this document, the DASD Component consists of numerous modules working together to (1) shape system performance and utilization data for detailed analysis by other modules, (2) evaluate the data to assess potential causes of performance, (3) describe the results from the evaluation, and (4) maintain a historical record of the results from the analysis. These modules are loaded and controlled by the central DASD Component of CPExpert (titled DASCPE).

Step 1. Use TSO ISPF to change the "prefix" in the data set names

Use TSO ISPF to change the "prefix" in the data set names (DSN) in the **USOURCE** DD statement, the **SOURCE** DD statement, the **CPEDATA** DD statement, the **CPEDASD** DD statement, the **HISTORY** DD statement, the **PDBLIB** DD statement, and the **SYSIN** DD statement of the JCL in accordance with your installation standards. The JCL is illustrated in Exhibit 4-1. (A "shell" of this JCL is contained on the distribution tape as "DASJCL2")

```
//jobname    JOB   job card information
//STEP02     EXEC  SAS,OPTIONS='MACRO DQUOTE PAGESIZE=65 ERRORABEND'
//USOURCE    DD    DSN=prefix.CPEXPERT.USOURCE,DISP=SHR
//SOURCE     DD    DSN=prefix.CPEXPERT.SOURCE,DISP=SHR
//CPEDATA    DD    DSN=prefix.CPEXPERT.CPEDATA,DISP=OLD
//CPEDASD    DD    DSN=prefix.CPEXPERT.CPEDATA,DISP=OLD
//PDBLIB     DD    DSN=prefix.MXG.MON,DISP=SHR
//LIBRARY    DD    DSN=saslib containing MXG FORMATs
//SYSIN      DD    DSN=prefix.CPEXPERT.SOURCE(DASCPE),DISP=SHR
```

JOB CONTROL LANGUAGE TO EXECUTE THE DASCPE MODULE

EXHIBIT 4-1

The CPEDATA DD statement in Exhibit 4-1 refers to the SAS data library maintained by CPExpert. The space for this library was created during the installation of CPExpert.

The CPEDASD DD statement in Exhibit 4-1 refers to the SAS library containing the DASD information created by the CPExpert modification to MXG or MICS. The CPEDASD DD statement is required only if you have installed the modification to MXG or MICS necessary to allow CPExpert to collect DASD I/O information at the job step level.

The PDBLIB DD statement in Exhibit 4-1 refers to the SAS library containing the performance data base to be analyzed. The example shows a sample DSN for a typical MXG performance data base. The DSN would be changed to "DSN=prefix.RMF.MICS.DETAIL" to use a MICS performance data base.

Exhibit 4-1 does not show the optional DD statements for MICS Information Areas (i.e., BATLIB DD, SCBLIB DD, HARLIB DD, etc.). The *CPExpert Installation Guide* describes how to use these optional DD statements if you have your MICS performance data base separated by MICS Information Area.

The LIBRARY DD statement in Exhibit 4-1 refers to the SAS library containing the MXG FORMATs. If you execute CPExpert against a MICS performance data base, the SASLIB DD statement would be changed to refer to the MICS FORMAT library. (CPExpert actually does not use the MICS FORMATS. You can specify the SAS **NOFMTERR** option to eliminate problems with SAS FORMAT errors.)

Step 2: Make any appropriate changes to the DASGUIDE Module

Before submitting the JCL shown in Exhibit 4-1 and executing the DASDCPE Module, you should make appropriate changes to the CPEXPERT.USOURCE(DASGUIDE) module. These changes are described in Section 3 of this manual.

Step 3. Execute the DASCPE Module

Submit the JCL to execute the DASCPE Module. Most installations execute the DASCPE Module on a daily basis, after their normal update of their performance data base.

Chapter 2: Executing the DAS1415 Module

This section describes how to execute the DAS1415 Module of the DASD Component. The DAS1415 Module is executed **only** if you wish to perform expanded analysis based on data set name.

The DAS1415 Module reads the SMF Type 14/15 records to acquire information necessary to identify the DD names associated with data set names defined in USOURCE(DASGUIDE). As explained in Section 1 (Chapter 5.3), the DD names are later used in a modification to MXG or MICS processing to identify the activity of devices associated with the DD names.

The DAS1415 Module must be executed before MXG or MICS are executed to perform their normal daily update of your performance data base. This is necessary because the modification to MXG or MICS reads the data set created by the DAS1415 Module. [The data set created by the DAS1415 Module is simply the output from a SAS PROC FREQ creating a distribution of unique DD names which were used to reference data sets defined in USOURCE(DASGUIDE).]

Step 1. Use TSO ISPF to change the DD statements

Use TSO ISPF to change the SMF DD statement to refer to the current SMF data. Use TSO ISPF to change the "prefix" in the USOURCE DD statement, the SOURCE DD statement, the CPEDASD DD statement, and the SYSIN DD statement of the JCL in accordance with your installation standards. The JCL is illustrated in Exhibit 4-2. (A "shell" of this JCL is contained on the distribution tape as "DASJCL3")

```
//jobname    JOB   job card information
//STEP1      EXEC  SAS,OPTIONS='MACRO DQUOTE PAGESIZE=65 ERRORABEND'
//SMF        DD    DSN=data set name of current SMF file,DISP=SHR
//USOURCE    DD    DSN=prefix.CPEXPERT.USOURCE,DISP=OLD
//SOURCE     DD    DSN=prefix.CPEXPERT.SOURCE,DISP=SHR
//LIBRARY    DD    DSN=saslib containing MXG FORMATs
//CPEDASD    DD    DSN=prefix.CPEXPERT.CPEDASD,DISP=OLD
//SYSIN      DD    DSN=prefix.CPEXPERT.SOURCE(DAS1415),DISP=SHR
```

JOB CONTROL LANGUAGE TO EXECUTE THE DAS1415 MODULE

EXHIBIT 4-2

Step 2. Execute the DAS1415 Module

The next step is to execute the DAS1415 Module to process the USOURCE(DASGUIDE) information and to process the SMF Type 14/15 data.

The DAS1415 Module will create two SAS data sets in the CPEDASD SAS library, containing information extracted from USOURCE(DASGUIDE) and containing information extracted from the SMF Type 14/15 records for data sets matching the data set names defined in USOURCE(DASGUIDE).

Checklist for Executing the DASD Component, Mainframe

☐ Execute the DASD Component.

  ☐ Change the "prefix" in the data set names in the DD statements.

  ☐ Make any necessary changes to the DASGUIDE Module in USOURCE.

  ☐ Submit the JCL to execute the DASDCPE Module.

Checklist for Executing DASD Component, Personal Computer

☐ Execute the DASD Component.

- ☐ Identify the directories to SAS and CPExpert.
  - ☐ USOURCE filename at SAS PGM window
  - ☐ SOURCE filename at SAS PGM window
  - ☐ LIBRARY filename at SAS PGM window

- ☐ Make any necessary changes to the DASGUIDE Module in USOURCE.

- ☐ If you are executing under Windows, enter "%INCLUDE SOURCE(DASCPE)" at SAS PGM window and submit.

- ☐ If you are executing under OS/2, enter "%INCLUDE SOURCE(DASCPE.SAS)" at SAS PGM window and submit.

Checklist for Performing Expanded Analysis

This checklist contains additional steps which must be performed if you wish the DASD Component to perform expanded analysis.

☐ Make sure that the CPExpert modification to MXG or to MICS is properly installed.

☐ Specify **%LET TYPE30DD = Y;** in USOURCE(GENGUIDE) to tell CPExpert that SMF Type 30(DD) information is available.

☐ If you wish CPExpert to analyze DASD performance from the perspective of critical workloads, take the following steps:

  ☐ Define workload categories in USOURCE(DASGUIDE).

  ☐ If you are running MVS Goal Mode), be sure that **%LET GOALMODE=Y;** has been specified in USOURCE(GENGUIDE).

  ☐ Specify **%LET LOVED1 = xxxx** in USOURCE(DASGUIDE), where 'xxxx' is the name of a "loved one" workload.

  ☐ Execute the DASCPE Module of the DASD Component.

☐ If you wish CPExpert to analyze DASD performance from the perspective of critical data sets, take the following steps **BEFORE** your performance data base has been updated:

  ☐ Specify critical data set names and their associated response objectives in USOURCE(DASGUIDE).

  ☐ Execute the DAS1415 Module to extract information from SMF Type 14/15 records.

☐ If you wish CPExpert to analyze DASD performance from the perspective of critical data sets, take the following steps **AFTER** your performance data base has been updated:

  ☐ Make any desired modifications to data set names and their response objectives in USOURCE(DASGUIDE).

  ☐ Specify **%LET DASDSN = Y** in USOURCE(DASGUIDE) to tell CPExpert to perform data set name analysis.

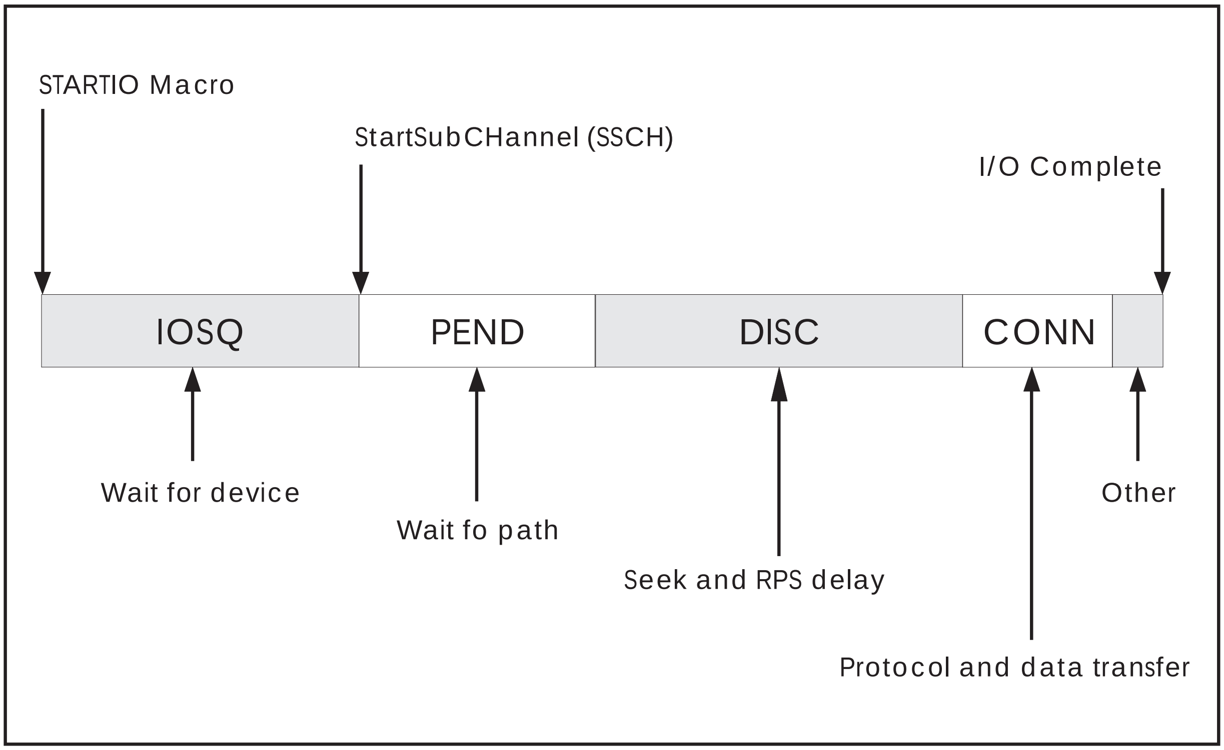
  ☐ Execute the DASCPE Module of the DASD Component.

Section 5: DASD Analysis Factors

This section discusses DASD performance analysis considerations. Chapter 1 presents an overview of performance factors from a DASD I/O operation viewpoint. Chapter 2 highlights some of the factors that must be considered when analyzing DASD performance based upon data collected and recorded by SMF or RMF.

Chapter 1: Overview of DASD Performance Considerations

From a high-level view, there are four key measures of DASD performance: IOS Queue (IOSQ) time, pending (PEND) time, disconnect (DISC) time, and connect (CONN) time. These measures are reported by RMF in SMF Type 74 records. Exhibit 5-1 illustrates these four measures and another potential element of DASD I/O time, titled "Other".



MAJOR COMPONENTS OF DASD I/O OPERATIONS

EXHIBIT 5-1

Chapter 1.1: IOSQ time

IOSQ time is the time from the issuance of a STARTIO macro until the Start SubChannel (SSCH) instruction is issued. After the STARTIO macro is issued, the software determines whether the device is busy. If the device is not busy with this system, the SSCH instruction is issued. However, if the device is busy with this system, the I/O request is queued.

Thus, IOSQ time always means that the device is unable to handle additional requests from this system. (The emphasis on "this system" is explained in the below discussion of PEND time.)

Some small IOSQ time is often unavoidable. However, large IOSQ time imply a situation that should be examined. Large IOSQ times result from (1) too many I/O operations directed to the device or (2) lengthy device response times (perhaps caused by high seeking, high RPS delays, or high PEND time). Large IOSQ times usually involve the following situations:

- Multiple data sets may be active on the volume. This situation is the most common and easiest to solve. The data sets can be redistributed among different volumes, to eliminate the queuing for the single volume.

- Multiple users may be using the same data set on the volume. Depending upon the data set characteristics, duplicate copies of the data set placed on different volumes may solve the IOSQ problems.

- Multiple application systems may be using the volume experiencing high IOSQ times. In this case, perhaps application redesign or scheduling can solve the problem.

- A particular application (or system function) may be executing I/O to the device faster than the device can respond.

- The overall device response time (PEND, DISC, and CONN) times may be large, such that the device is unable to provide quick response to the I/O requests. This situation will be revealed by large values in the PEND, DISC, or CONN measures.

With Parallel Access Volumes (PAV) and dynamic alias management in Goal Mode, IOSQ time should be significantly reduced or eliminated. The implications of PAV and dynamic alias management will be discussed in rules related to these features.

Chapter 1.2: PEND time

PEND time is the time from the issuance of the SSCH instruction until the device is selected by the control unit. This time is caused by queuing for the path (wait for channel, wait for control unit or wait for head-of-string), and can be caused by other systems sharing the device (wait for device). Large PEND times usually involve the following situations:

- **Shared devices**. If the device is shared with another system, PEND time may indicate contention with the other system. Large PEND times in shared-device environments usually involve situations very similar to those described under IOSQ time:

  - Multiple data sets may be active on the volume. This situation is the most common and easiest to solve. The data sets can be redistributed among different volumes, to eliminate the queuing at the channel level (reflected as PEND time) for the single volume.

    If some of the data sets are not required to be shared, then the Data Base Administrator has complete flexibility to move these data sets (subject, of course, to the performance implications of the target devices). These data sets should be moved to a non-shared device.

    If the data sets are required to be shared, then they must be relocated to shared devices.

  - Multiple applications or users may be using the same data set on the volume. Depending upon the data set characteristics, duplicate copies of the data set may be placed on different volumes. This would solve the PEND problems cause by contending systems. If this option is feasible, the data sets could be placed on non-shared devices, likely resulting in even more performance improvement.

  - Multiple application systems may be using the volume experiencing high PEND times. In this case, perhaps application redesign or scheduling can solve the problem.

  Additionally, large PEND times for shared devices could be caused by RESERVE from the other system. The applications issuing the RESERVE should be examined to determine whether the RESERVE is required. If the RESERVE **is** required, the above situations should reviewed to determine whether improvements can be achieved.

- **Non-shared devices**. Large PEND times for devices that are not shared may mean that there are insufficient paths available to the device. Too much I/O may be directed to many devices on the path, control unit, or head-of-string. The data sets can be redistributed among different volumes on different paths, control units, or heads-of-string. This will reduce the hardware-level queuing. Alternatively, the entire volume may be moved to a different (less busy) head-of-string or path.

  If redistributing the data sets or moving the volume is not feasible, then the device should have more paths. Depending upon the existing configuration, this may involve re-configuring existing channel paths, or acquiring additional hardware.

Fortunately, SMF Type 78 and Type 74 records contain information which can be used to identify at which level the hardware queuing occurs (that is, whether queuing is for the path, control unit, or head-of-string).

- **Devices attached to cached controllers**. Large PEND times for devices attached to cached controllers may imply a high percent of read miss operations, or non-volatile storage (NVS) writes for IBM-3990-3 devices. Fairchild [1] lists four ways in which staging in caching controllers can cause hidden device busy (with the device busy potentially reflected in high PEND time):
  - The normal (random) caching algorithm stages all records to the end of the track after a requested record is read.
  - The normal (random) caching algorithm stages all records from the beginning of the track to the requested record if a front-end miss occurs.
  - The sequential caching algorithm stages all records to the end of the track after the requested record is read, and stages in all of the next track. IBM-3990 (Model 3) controllers stages in all of the next three tracks.
  - Most writes to extended function IBM-3990 (Model 3) go into NVS with a subsequent destaging required.
- **Dual Copy Initialize**. Large PEND times for IBM-3390 devices may be caused by dual copy initialize. In this case, the dual copy initialize should be turned off.

Chapter 1.3: DISC time

DISC time is the time (1) from when the controller initiates a SEEK Channel Command Word (and the seek requires an arm movement) on the device until the SEEK command is complete, (2) plus the time of the rotational delay while the SET SECTOR Channel Command Word is executing, and (3) plus the rotational position sensing (RPS) delay time required because of missed RPS reconnect.

- **Seek delay**. The SEEK command is responsible for positioning the arm to the proper cylinder. If no positioning is required (that is, the arm is already at the proper cylinder), the device is not disconnected. Seek operations occur because of accessing patterns with data sets and because of accessing patterns between data sets. Large seek times usually involve the following situations:

[1] Fairchild, Bill, "The Anatomy of an I/O Request", *Conference Proceedings*, CMG'90, the Computer Measurement Group, Chicago, IL.

- Multiple data sets being active on the volume. The data sets can be redistributed among different volumes, to eliminate the seeking on the single volume.

- Multiple users using the same data set on the volume. While only one data set is involved, the user or application accessing patterns may require frequent arm movement. A partitioned data set in which several TSO users reference different members is a common situation.

  Depending upon the data set characteristics, duplicate copies of the data set placed on different volumes may solve the seeking problems.

- **Rotational delay**. The SET SECTOR command is responsible for locating the proper sector on the track as the disk rotates. (Actually, the SET SECTOR command locates a sector three sectors preceding the desired sector. This sector is called the *angular sector.*) The device is disconnected during the SET SECTOR command operation.

  The rotational delay may be from zero, to the total time required to rotate the disk to the required sector. It is possible that the required sector will be immediately under the head. In this case there is zero rotational delay. On the other hand, the sector could have just passed under the head before the SET SECTOR command was received by the drive. In this case, a full rotation must be accomplished before the required sector is located. On average, one-half of the rotation time will be required to locate the sector. This time is referred to as the **average latency** of the device. For example, IBM-3380 devices rotate every 16.6 milliseconds and the average latency is 8.3 milliseconds.

  It is important to realize that the latency is an average based upon many SET SECTOR commands. Any particular SET SECTOR command may have a latency ranging from zero to the maximum rotational delay.

  If there are few I/O commands for a particular device in a given measurement interval, it is uncertain what the average latency will be. However, if there are many I/O commands for a particular device in a given measurement interval, the average latency will normally be one-half of the rotational delay.

  The average latency may be (and should be) quite small with cached devices. This is because many I/O requests should be satisfied from the cache and have no latency.

- **Missed RPS reconnect**. The device attempts to reconnect to the path when the angular sector is reached (the angular sector is described above). If the reconnect attempt is successful before the desired sector is reached, then the device connects and the read or write operation can proceed. If the reconnect is not successful before the desired sector is reached, then the device does not connect, and a complete revolution of the track must occur before the angular sector is again reached. This is called a *missed rotational position sensing reconnect (or missed RPS reconnect)* delay.

There is no action which can alleviate the initial rotational positioning delay (aside from changing device characteristics, such as implementing caching or buffering at the device level). Over a large number of I/O operations, this initial delay will be one-half the rotation times.

However, the missed RPS delay is a function of the probability that the path will be busy when the device attempts to reconnect; the busier the path(s), the more missed RPS delay. (Note that the path busy time is a function of the connect time of other actuators. The path cannot be busy from the device itself when the device attempts to reconnect.)

There are no measurements available to determine precisely how the DISC time should be divided. However, queuing models can estimate the RPS reconnect delay time and assumptions can be made about the average latency time. By subtracting RPS reconnect delay and the average latency time from the DISC time, the result will be an approximation of the time spent seeking. The results from this method is valid if (1) the distribution of device access is random so that the queuing formulae are applicable and (2) a large number of I/O operations are involved so that the average latency assumption is proper.

Chapter 1.4: CONN time

CONN time is the time in which the device is actually connected to the path. This time includes the data transfer time, but also includes protocol exchange[2] (or "hand shaking") between the various components at several stages of the I/O operation.

The data transfer time obviously is a function of the amount of data being transferred. This simply is the number of bytes transferred divided by the transfer speed (for example, if 4096 bytes were transferred from an IBM-3380 with a transfer speed of 3,000,000 bytes per second, the 4096 bytes would require 4096/3,000,000 seconds; or about 1.36 milliseconds).

Large connect times generally are caused by the following situations:

- A large average block size. This situation may be highly desirable for sequential data sets, but would be quite undesirable for randomly accessed data.

- Long multi-track searches. For example, the catalog must be searched for cataloged files, the Volume Table of Contents (VTOC) must searched to find a requested file, a directory must be searched for partitioned data sets, etc.. These searches will result in long connect times for the volume involved.

[2]Note that the protocol exchange occurs at multiple points in the normal I/O operation, even though it is shown only once in Exhibit 5-1.

- Fairchild indicates that wrong sector numbers, non-IBM temporary error recovery, and coding OPTCD=W can also cause long connect times.

There have been no published "official" times for the protocol connect times. The protocol time has been estimated by different authors to range from .5 milliseconds to 1.5 milliseconds, depending upon the device and the authors' experiences. In any event, there generally is little that can be done about the protocol time.

Chapter 1.5: OTHER time

There are at least two other potential I/O delays for DASD: (1) waiting for the I/O completion interrupt to be serviced by a processor and (2) waiting for the I/O interrupt to be serviced by a domain under PR/SM. Neither potential I/O delay is expected to be of the magnitude of the four "standard" I/O delays. However, they can be significant in special circumstances.

- Multi-processor configurations running under MVS can use any processor to service an I/O interrupt. However, when a processor services an I/O interrupt, the processor's high-speed cache storage is no longer valid when control is returned to the interrupted task. Consequently, many of the processor's high-performance design features may be nullified.

  A hardware feature allows processors to be disabled for I/O interrupts. With this method, only a small number (perhaps only one) processor is enabled for interrupt processing. Only this processor will have its high-speed cache storage disturbed by the task-switching required for interrupt processing, and only this processor will periodically have its high-performance design features nullified. The disadvantage to this approach is that an interrupt may occur while the processor is busy servicing a previous interrupt.

  If an interrupt is pending and no processor is enabled to service the interrupt, the interrupt must wait until a processor is available. This time should be insignificant, unless the system is processing a significantly large number of I/O operations. If the system is processing a large number of I/O operations, the interrupt pending delay could pose performance problems.

  After the processor completes processing for an I/O interrupt, it issues a Test Pending Interrupt (TPI) instruction to determine whether there are any interrupts pending. If an I/O interrupt is pending, the processor proceeds to service that interrupt.

  The IEAOPTxx member of SYS1.PARMLIB contains the **CPENABLE** keyword. This keyword specifies the percent of I/O interrupts detected by the TPI instruction, compared with all I/O interrupts. When the percent exceeds the high threshold of the CPENABLE keyword, MVS enables another processor to handle pending I/O interrupts. If the percent falls below the low threshold of the CPENABLE keyword, MVS will disable a processor (to the point that only one processor is enabled). The low and high threshold values for CPENABLE are 10 and 30 percent, respectively. These values normally mean that less than 30% of the I/O interrupts will be delayed for I/O interrupt service.

- MVS environments running under as a guest under VM or in a logical partition (LPAR) under PR/SM are subject to I/O interrupt delays. These delays can occur if another guest (for VM) or another domain is in its dispatch interval when the I/O interrupt completion is posted. The I/O interrupt remains pending until the guest or domain is dispatched. These delays have been estimated to be far more significant than might otherwise be expected.

Neither of the potential I/O delays described above is measured by RMF (although RMF does provide information on the number of I/O interrupts serviced by each processor and the number of TPI instructions resulting in I/O interrupt servicing).
The potential I/O delays are included in this discussion of general DASD performance considerations because (1) they may become important under certain situations and (2) techniques may be developed to assess their impact.

Chapter 2: RMF Data Analysis Considerations

This chapter highlights some of the factors that must be considered when analyzing DASD information collected and recorded by SMF in Type 30 records or by RMF in Type 70(series) records.

These factors do **not** preclude a comprehensive analysis of performance data and usually do not prevent insight into the causes of unacceptable performance. However, the factors must be recognized and accounted for both by CPExpert in analyzing data and by the user in reviewing CPExpert's results. The factors stem from (1) the way in which SMF and RMF create and record Type 30 and Type 70(series) information, and (2) inherent limitations caused by data averages.

Chapter 2.1: SMF information

SMF Type 30 records contain a record sub-type code to identify when the records are written:

CODE	SUB-TYPE DESCRIPTION
1	Job start
2	Interval records
3	Step termination
4	Step total
5	Job termination
6	System address space

SMF will optionally record the different sub-types, depending upon parameters contained in the SMFPRMxx member of SYS1.PARMLIB. Most installations collect Sub-type 4 (Step total) records, and many installations collect Sub-type 2 (Interval) records. If Interval records are recorded by SMF, Sub-type 3 (Step termination) records are automatically created.

It is highly desirable to collect Interval/Step termination information. It is virtually impossible to analyze system performance based upon Step total information if there exists long-running jobs. This is because it is impossible to correlate the information reflected in the Step total records with the information contained in SMF Type 70(series) data.

The Sub-type 4 records are written only after a long-running job step terminates, while the SMF Type 70(series) records are written at user-defined intervals (the interval typically is every 30 minutes or so). Long-running job steps may span many RMF recording intervals [RMF is responsible for creating the SMF Type 70(series) records]. Consequently, there may be many RMF interval records written between the start and end of a long-running job step.

Sub-type 2 (Interval) records are written at user-defined intervals (typically the interval selected is the same interval as the RMF interval records). SMF writes a Sub-type 2 record when the specified interval has lapsed after the start of the job step and continues to write Sub-type 2 records at each subsequent interval. When the job step terminates, SMF writes a Sub-type 3 record containing the information since the last Sub-type 2 record was written. One consequence of the interval records is that system usage can be identified by workload, and can be correlated with the overall system statistics recorded by RMF in the SMF Type 70(series) records.

There are two variations in how SMF and RMF write interval data: (1) non-synchronized and (2) synchronized. Synchronization of SMF and RMF records is an option that must be explicitly specified in the SMFPRMxx member of SYS1.PARMLIB.

- **Non-synchronized writing of SMF and RMF interval data.** With non-synchronized writing of SMF and RMF interval data, the Sub-type 2 records are written based upon the interval lapse from the start of the job step. They are not written at the same time as is the SMF Type 70(Series) records. This lack of coordination between recording the two record types poses a correlation problem: a particular Sub-type 2 (or Sub-type 3) record may span between two RMF recording intervals. From a data analysis view, there is no way to precisely determine whether the data reflected in the Sub-type 2 record (or Sub-type 3 record) should belong to the first RMF Type 70(series) interval or should belong to the second RMF Type 70(series) interval.

  For example, suppose that the RMF recording interval were specified as 30 minutes, and RMF was directed to synchronize on the hour and half-hour. The RMF data would be collected and recorded at 10:00, 10:30, 11:00, 11:30, etc. Further suppose that a particular job step started at 15 minutes past the hour. Assuming that the Type 30 interval recording were specified as 30 minutes, SMF would create a Type 30 (Sub-type 2) interval record at 45 minutes past the hour, 15 past the next hour, and so forth. Thus, the RMF data would be recorded on the hour and half-hour, while the Sub-type 2 data would be recorded "offset" by 15 minutes.

  CPExpert addresses this problem by pro-rating the SMF Type 30 information based upon elapsed time. In the above example, 50% of the actual workload data contained in the SMF Type 30 (Sub-type 2 or Sub-type 3) records would be attributed to one RMF measurement interval and 50% would be attributed to the next RMF measurement interval. This pro-rating approach essentially assumes that the resources required by a job step do not vary much from one instant to the next.

  This approach works quite nicely so long as the job step uses resources in a uniform fashion. Many job steps exhibit this characteristic, and the resources required by the job step do not vary much as the job executes. Resources distributed using the pro-rating approach result in fairly consistent usage characteristics when comparing the summarized Type 30 data with SMF Type 70(series) data.

However, some job steps exhibit significant cycles, or require resources at the beginning or end of the job step. For these job steps, the pro-rating approach does not properly distribute the resource usage into the correct RMF measurement interval. Summarized Type 30 data would not compare well with Type 70(series) data if many job steps exhibit this cyclic or burst nature of resource usage. Unfortunately, there is no way to better distribute the data. Consequently, analysis based upon Type 30 data must be viewed with some caution. The analysis **generally** will be sufficiently precise for performance analysis purposes. However, anomalies will appear and results must always be subjected to a "reality" test.

This point is significant for the DASD Component, because the DASD Component attributes DASD usage to workloads based upon correlating SMF Type 30 data with SMF Type 74 data. The DASD I/O activity at the job step level is obtained from the SMF Type 30 interval records (using a modification to MXG or to MICS). This DASD I/O activity is pro-rated to the RMF measurement intervals as described above. The RMF DASD device I/O characteristics (IOSQ, PEND, DISC, and CONN times) are attributed to workloads based upon the pro-rating.

It is possible that the pro-rating method will result in improper attribution of I/O device characteristics to workloads. For example, suppose that a job step completed a few minutes past the hour (and that RMF data records were synchronized on the hour and half-hour). When the job step completed, it could execute many DASD I/O operations. These I/O operations would mostly be attributed to the previous RMF interval and the DASD device characteristics of that interval would be associated with most of the I/O operations. Suppose that there were no I/O problems with the device in the first interval. It is possible, however, that as the job step completed, it experienced significant DASD I/O problems which would be reflected in the second RMF interval. Since only a few of its I/O operations would be attributed to the second RMF interval, CPExpert would associate the I/O problems to only a few of its I/O operations. Consequently, CPExpert might consider the job step (and the workload category associated with the job step) to receive good DASD service, when the workload actually received bad service because of its burst I/O operations.

This example is not intended to invalidate the techniques CPExpert uses. Rather, the example is presented to explain a unique situation in which the techniques could result in improper conclusions. Readers may note that IBM's Service Level Reporter and any other software analyzing SMF/RMF data face the same analysis problem. The problem is with the data; not with the technique.

- **Synchronized writing of SMF and RMF interval data.** Synchronized writing of SMF and RMF was introduced with MVS/ESA SP4. When interval accounting is synchronized, SMF generates interval records for a work unit based on the end of the

SMF global recording interval, rather than the start time of a job. This feature allows Type 30 records (and other record types) to be synchronized with writing RMF Type70(series) records. SMF places indicators (or “flags”) in SMF Type70(series) records to indicate whether SMF and RMF records are synchronized.

It is not necessary for CPExpert to pro-rate data if the recording intervals are synchronized.

Chapter 2.2: Data Averages

The data collected by RMF and recorded in SMF Type 70(series) records provide a valuable source of information about the use and interaction of system components and workloads. However, the data are summarized and recorded at specific intervals (e.g., every 30 minutes). For most data elements, analysis must be accomplished based upon the **summary** or **average** values.

For example, the DASD IOSQ time reported for each device is the total for the measurement interval. The average IOSQ time per I/O operation is computed by dividing the total IOSQ time by the number of I/O operations. This average may have no relation to the IOSQ time experienced by any particular I/O operation. This problem is particularly pervasive as the RMF recording interval becomes more lengthy (e.g., if the recording interval were 60 minutes). The DASD IOSQ time may be quite long during the first half of the RMF measurement interval when contending workloads execute. The DASD IOSQ time may be short during the last half when a workload executes without contention from other applications. The average of the two extremes may lead to a conclusion that there was no problem with DASD IOSQ time for the entire interval!

Most DASD analysis performed by CPExpert assumes either a uniform or an exponential distribution of DASD I/O operations. For example, the pro-rating discussed in the previous chapter assumes a uniform distribution of I/O operations on a **job step** basis, over the life of the job step. However, the queuing models employed by CPExpert to analyze various aspects of DASD delays generally assume an exponential distribution of I/O operations at the **device** level, over an entire RMF measurement interval. Neither of these assumptions may be correct.

Wicks[3] illustrates a variety of distributions of the arrival rate of events, ranging from uniform distribution, to "cafeteria" distribution (events mostly arrive in clusters), to "London bus" distribution (events arrive only in clusters), to a random distribution (events exhibit a Poison or exponential arrival). The arrival of many events in computer systems exhibit an exponential distribution, and M/M/1 or M/M/C queuing models can fairly represent many aspects of the systems.

[3] Wicks, R. J., "*Balanced Systems and Capacity Planning*, IBM Corporation Washington Systems Center Technical Bulletin GG22-9299-02

However, Wicks gives an excellent example of exceptions: when editing a dataset using ISPF, the entire dataset may be read, some time is spent editing, and then the entire dataset may be written. The I/O requests in this instance would be similar to Wicks' "London Bus" distribution.

There is a tradeoff between (1) recording RMF data frequently, incurring the overhead and storage requirements of the additional RMF records, and requiring additional resources to analyze the data versus (2) recording RMF data less frequently, having less precise or representative data to analyze, and minimizing the resources required to perform the analysis. The importance of these tradeoffs must be evaluated in light of the objectives of the analysis and the requirements for precision of results.

In any event, any review of analysis and conclusions (whether by CPExpert or by a performance analyst) must be viewed with some caution because of the data summary, data averaging, and data distribution issues.

If the analysis consistently results in the same conclusions, you can be reasonably sure that the analysis is correct. However, it generally is unwise to make changes based upon analysis of a single day's RMF measurement information unless a "reality test" indicates that the analysis clearly is correct.

Section 6: Using the DASD Component

This section describes how to use the DASD Component to analyze system performance in the areas of DASD performance evaluated by the DASD Component.

As discussed in Section 1, the DASD Component is designed based upon the DASD performance analysis concept of "address the outrageous problems" rather than attempting to identify every potential problem. Most Data Base Administrators will be quite pleased to solve one major DASD problem at a time. Consequently, using the DASD Component involves (1) executing the software, (2) reviewing the output, (3) making system changes or altering the guidance to the DASD Component, and (4) iterating through the process.

This process is described in terms of providing guidance to the DASD Component, actions that should be taken on a daily basis (or more frequently if you wish), and actions that should be taken on a weekly or monthly basis.

Chapter 1: Prepare guidance for the DASD Component

Use TSO ISPF to change the CPEXPERT.USOURCE(DASGUIDE) PDS member to reflect the guidance required by the DASD Component. Section 3 describes the variables that must be changed in the DASGUIDE PDS member. This step must be taken only when the guidance changes (normally when you wish to select different to analyze, when you wish to exclude additional volumes, or when you wish to redefine workloads).

Chapter 2: Actions on a daily basis

Use the DASD Component to analyze overall system performance, by following the below steps:

Step 1: Execute the DASCPE Module

The JCL to execute the DASCPE Module is described in Exhibit 4-1. We suggest that you append the JCL to execute the DASCPE Module to the normal daily update of your performance data base.

Step 2: Review the output from the DASCPE Module

If any rules were produced, refer to the specific rule in Appendix A for a description of the rule, a discussion of why the rule was produced, and a recommendation for actions that should be taken.

Depending upon the output, you may wish to make changes or wait to see if the problems are identified in an analysis of a subsequent day's data.

- The DASD Component may identify problems which clearly should be solved because their effect is so serious. In many cases, once the problem is identified, users immediately realize that the problem and suggested solutions make sense.

- The DASD Component may identify problems which you do not feel will commonly occur. For example, you may review the "worst performing device" and realize that it is the "worst" only because of an infrequent application scheduling situation. In this case, you may wish to exclude the volume and reprocess the data. (The DASD Component will provide summary information about the next "worst" devices, so you can appreciate whether it is likely that there are other serious problems.)

- The DASD Component may identify problems about which you have doubts. The discussion in Section 5 illustrates the potential problems with the analysis, and the discussion associated with each rule often amplifies the cautions appropriate to the particular analysis. In this case, you probably should take no action but wait to see whether the same problems and analysis recur with subsequent data. If the same problems are identified by the DASD Component after analyzing several day's SMF/RMF data, you can be more confident that SMF/RMF data recording problems play less of a consideration.

Chapter 3: Actions on a weekly or monthly basis

Review the recommendations produced during the previous week or month. If the same problems consistently occur and CPExpert makes consistent recommendations, then you should consider action. This deliberate review of the problems and associated recommendations ensures that the problems are continuing and that change is warranted.

You generally should make only one change at a time! This sound tuning advice if founded on the principles that:

- Tuning is an art. No one (and certainly not CPExpert) can guarantee that any particular change will have a beneficial effect in all environments.

- Changes may have unexpected effects. Most systems are complex, parameters may improve performance of one area at the expense of performance in another area, and management may wish resources focused on the second area.

- If you make multiple changes and performance deteriorates, you will be unable to identify easily the change causing the problem. You are then faced with the problem of backing out all of the changes and starting over, one at a time.

- Some changes are not "precise" in that, for example, keyword values might need to be adjusted a little at a time until a suitable value is reached. If multiple changes are made, you will be unable to detect the effect of the fine- tuning of the changes.

Above all, **remember that the recommendations from CPExpert are simply options** to be considered in the context of overall objectives. You must decide whether the recommendations are reasonable. Rarely should a recommendation be implemented without first evaluating how the recommendation will effect other workloads.

Please remember that CPExpert is not intended to replace a performance analyst. Rather, CPExpert was developed to <u>help</u> analyze the performance of MVS systems. CPExpert automates much of the routine of computer performance evaluation. Performance analysts can then focus on the areas which are not routine and which "require thinking".

With this philosophy, please let us know when you discover areas in which CPExpert could have helped you analyze a problem. We will improve our product and you will have more help!

This request is particularly applicable to the DASD Component. We could have implemented much additional analysis and many additional options. We decided to wait to see what our users **wanted** and respond to your requirements, rather than trying to "guess" about your DASD analysis needs. However, we are eager to provide additional analysis; let us know what you want!

Description of Rules

Appendix A

This appendix contains a description of each rule that results in a finding by the DASD Component of CPExpert. The description summarizes the rule, lists predecessor rules, discusses the rationale for the finding, and recommends action.

The summary of the rule presents a short description of the finding.

The predecessor rules are listed so you can follow the line of reasoning leading to a particular rule being executed.

The discussion describes as much as necessary of the operation of the computer system (the hardware, MVS, the Workload Manager, etc.) as it relates to the particular rule. The purpose of the discussion is to explain the reasoning behind the rule, and what causes the rule to be produced.

The recommendations suggest possible actions that should be considered based on the findings. In many cases, multiple possible actions are listed. You must determine which actions should be taken (this determination is based upon the suitability of the actions to your own environment, the financial implications of the action, and the "political" acceptability of the action.)

The rules are organized in numerical order. However, not all numbers are represented. The LIST OF RULES in this appendix lists all rules that are included in the initial release of the DASD Component.

The DAS2nn rules are very similar to the DAS1nn rules. The major difference in the DAS2nn rules is that they relate to the "expanded" analysis of "loved one" workload. Consequently, the narrative is somewhat different. However, the DAS2nn rules will often refer you to the DAS1nn rules for more detailed discussion and suggestions.

You may wish to read all of the rules in this appendix, just to see the type of problems that are encountered in different installations. However, it is not necessary to read all of the rules. It is necessary only to read the rules that apply to your installation. The rules that apply to your installation are identified by the report produced from the DASCPE Module.

List of Rules

RULE	DESCRIPTION
DAS000	Sysplex performance characteristics of significant volumes
DAS050	Performance characteristics of significant volumes
DAS100	Volume with worst overall performance
DAS102	Volume with next worst overall performance
DAS105	Volume performance was not consistently poor in any area
DAS110	Seeking was the major cause of I/O response delay
DAS111	Seeking was probably caused by independent applications
DAS112	Seeking was probably caused by a single application
DAS113	Worst seeking was probably caused by independent applications
DAS114	Worst seeking was probably caused by a single application
DAS115	Seeking was cause of I/O delay on page pack
DAS120	Missed rotational position sensing was major cause of I/O response delay
DAS121	Volumes contributing to missed rotational position sensing
DAS123	Non-DASD devices contributed to RPS delay
DAS125	Applications contributing to RPS delay
DAS130	Large PEND time was major cause of I/0 response delay
DAS131	Large PEND time was caused by utilization of paths to device
DAS132	Large PEND time was caused by another system

 Revised: October, 2002

RULE	DESCRIPTION
DAS133	Large PEND time was caused by controller busy delays
DAS140	High connect time was major cause of I/O response delays
DAS150	High IOSQ was major cause of I/O response delays
DAS151	Queuing was probably caused by independent applications
DAS152	Queuing was probably caused by a single application
DAS153	Worst queuing was probably caused by independent applications
DAS154	Worst queuing was probably caused by a single application
DAS160	Missed cache read hits was major cause of response delay
DAS170	There did not appear to be a problem with the device
DAS180	Applications accessing the volume with the worst performance
DAS185	Applications accessing the volume during the period with worst performance
DAS200	Volume with worst overall performance from the perspective of the critical workload
DAS202	Volume with next worst overall performance from the perspective of the critical workload
DAS205	Volume performance was not consistently poor in any area
DAS210	Seeking was the major cause of response delay to the critical applications
DAS220	Missed rotational position sensing was major cause of response delay to the critical applications
DAS221	Volumes contributing to missed rotational position sensing

RULE	DESCRIPTION
DAS223	Non-DASD devices contributed to RPS delay
DAS225	Applications contributing to RPS delay DAS230 Large PEND time was major cause of response delay to the critical applications
DAS231	Large PEND time was caused by utilization of paths to device
DAS232	Large PEND time was caused by another system
DAS233	Large PEND time was caused by controller busy delays
DAS240	High connect time was major cause of response delays to the critical applications
DAS250	High IOSQ was major cause of response delays to the critical applications
DAS260	Missed cache read hits was major cause of response delay to the critical applications
DAS270	There did not appear to be a problem with the device
DAS280	Non-critical job steps used this volume and were a major cause of response delays to the critical applications
DAS285	Non-critical job steps used this volume during the period of worst performance and were a major cause of response delays to the critical applications
DAS287	Other applications did not reference the volume

RULE	DESCRIPTION
DAS300	Perhaps shared DASD caused performance problems
DAS310	Seeking probably was caused by shared DASD conflicts
DAS330	Large PEND time probably was caused by shared DASD conflicts
DAS360	Missed cache hits probably were caused by shared DASD conflicts
DAS380	Applications potentially causing shared DASD conflicts
DAS385	Applications potentially causing worst shared DASD conflicts
DAS390	Performance problems apparently were not caused by shared DASD conflicts
DAS400	Access characteristics of significant data sets

RULE	DESCRIPTION
DAS600	Excessive Control Area (CA) splits occurred
DAS604	Excessive secondary extents were allocated
DAS605	Excessive extents were used and secondary allocation unit was small
DAS606	Primary or Secondary allocation unit was small
DAS607	VSAM data set is close to maximum number of extents
DAS610	Relatively small CI size was used for sequential processing
DAS611	Relatively large CI size was used for direct processing
DAS612	Relatively large CI size was used for mixed processing
DAS620	The number of data buffers should be increased
DAS621	The number of index buffers should be equal to index levels
DAS622	The number of index buffers should be more than STRNO value
DAS625	NSR was used, but a large percent of the access was direct
DAS635	LSR was used, but a large percent of the access was sequential

Rule DAS000: Sysplex performance characteristics of significant volumes

Finding: CPExpert identifies the performance characteristics of the volumes in the sysplex that have the most potential for performance improvement.

Impact: This finding is used to assess the importance of the "worst" performing device and to determine whether other devices offer significant performance improvement potential.

Logic flow: This is a basic finding. There are no predecessor rules.

Discussion: CPExpert uses the following algorithm to identify the devices which have the most potential for improvement:

- CPExpert computes the average device response time for each **type** of device in the configuration, for each RMF measurement interval. The logic computes the average device response by type of device, since better performance would be expected from cached devices (for example) than from non-cached devices. This method essentially assesses the performance of each device against the performance of similar devices in the configuration.

- Devices that exceed the average device response time for their device type in any RMF measurement interval are selected as candidates for improvement. The rationale is that improvement efforts should not be directed at devices which provide better than average response. Thus, the candidate set of devices to analyze consists of those which provided worse than average response.

- The I/O rate of each “candidate device is weighted by its response time, **for the entire set of RMF intervals in which the device exceeded the average response**. The result is a measure of the relative performance improvement **potential** of each device which provided worse than average response, from an overall system view. For example, consider two devices in a device type having an average I/O response of 20 milliseconds:

  Device A: I/O rate = 30 I/O operations per second
  Device response = 25 milliseconds
  RMF intervals with above average response = 4
  Seconds per RMF interval = 900
  Weighting factor = 30 * 25 * 4 * 900 = 27,000,000

Device B: I/O rate = 5 I/O operations per second
Device response = 40 milliseconds
RMF intervals with above average response = 5
Seconds per RMF interval = 900
Weighting factor = 5 * 40 = 900,000

In the above example, CPExpert would select Device A as having the most overall potential for improvement, even though its per-I/O device response was not as bad as the device response of Device B.

CPExpert ranks the devices based on the weighting factor computed above. CPExpert then analyzes the devices, starting at the device with the highest weighting factor.

With Rule DAS000, CPExpert lists basic characteristics of the volumes having the most potential for improvement, so that you can appreciate the relative performance improvement potential between volumes on the list. The data presented by Rule DAS000 reflects the average per-second delays **only** during measurement intervals when the device I/O performance was worse than the average for its device type. This information is presented on a sysplex view basis, regardless of whether a specific system has been selected for analysis.

If the performance data base contains data for more than one sysplex, and if %LET SYSPLEX=*ALL; has been specified in USOURCE(DASGUIDE), CPExpert will produce information for all volumes in the performance data base. If a specific sysplex is selected for analysis (using the %LET SYSPLEX=xxxx, where “xxxx” is the name of a sysplex, only volumes for the designated sysplex will be listed.

The "weighted delays" value is a relative measure of the performance improvement potential of the volume. The absolute values in the column are not particularly meaningful. Rather, the values should be compared to each other to assess the relative performance impact of each volume.

It is possible that a volume may have a significant improvement potential in a particular measurement interval, but not be the volume with the most overall potential for improvement. This situation can arise because the analysis is directed toward the volumes with the **most overall** performance improvement potential. If you suspect that this is the case with a particular device, you can “select” that device for analysis, using the select process described in Section 3 of this document.

The following example illustrates the output from Rule DAS000:

```
RULE DAS000:  SYSPLEX PERFORMANCE CHARACTERISTICS OF SIGNIFICANT VOLUMES

   The following is a list of the most significant volumes showing their
   overall performance characteristics for the period being analyzed, from
   an overall sysplex view.  The "average per second delays" represent the
   averages ONLY during measurement intervals when the device performance
   was worse than the average for this device type on at least one system
   in the sysplex.  The "weighted delays" value is a measure of the overall
   relative performance impact of each volume.

                 DEVICE   I/O    ---AVERAGE PER SECOND DELAYS---- WEIGHTED
   SYSTEM VOLSER NUMBER  RATE    RESP   CONN   DISC   PEND   IOSQ  DELAYS
   SYZ0   SP0006  FE58  114.0 118.340  0.130  0.003  0.785117.423   236679
   SYH0   SP0006  FE58   93.2  94.382  0.106  0.003  0.775 93.498   188763
   SYF0   SP0006  FE58   40.7   5.983  0.048  0.001  0.510  5.425    11967
   SYE0   SP0006  FE58   28.5   3.444  0.035  0.001  0.404  3.005     6889
   SY90   DB008D  FDFA    2.9   0.755  0.035  0.028  0.007  0.685     1509
   SYA0   CAT00F  FD6B    1.5   0.413  0.018  0.001  0.006  0.387      826
   SYF0   D83IA1  BDE1    7.0   0.319  0.195  0.000  0.024  0.099      638
   SYA0   SP3057  FEC1    8.2   0.306  0.131  0.002  0.026  0.146      611
   SYE0   CAT00F  FD6B    1.1   0.299  0.013  0.001  0.005  0.281      599
   SYA0   DB0053  FEF9    1.4   0.298  0.014  0.003  0.004  0.277      597
   SYA0   CAT011  FD7C    6.1   0.279  0.014  0.090  0.058  0.116      558
   SYE0   CAT011  FD7C    5.3   0.256  0.013  0.065  0.061  0.117      512
   SY80   CAT00F  FD6B    1.0   0.245  0.012  0.000  0.003  0.229      490
```

In this example, it is clear that SP0006 has significant performance improvement potential. The DASD Component would analyze SP0006 from the view of each system in which it appeared as the “worst” device, to determine what caused the delays. Additionally, if the CPExpert modification to MXG or MICS (described in Section 2) had been installed, the DASD Component would list the applications referencing SP0006. Further, if SMF Type 42 records were available (and the volume contained data sets managed by DFSMS), the DASD Component would produce Rule DAS400 to show access characteristics of the most significant data sets that resided on SP0006.

Suggestion: You should use the information displayed by Rule DAS000 to assess the relative importance of the "worst" performing device compared with the performance improvement potential of the other devices.

Rule DAS050: Performance characteristics of significant volumes

Finding: CPExpert identifies the performance characteristics of the volumes in a system that have the most potential for performance improvement.

Impact: This finding is used to assess the importance of the "worst" performing device and to determine whether other devices offer significant performance improvement potential.

Logic flow: This is a basic finding. There are no predecessor rules.

Discussion: CPExpert uses the following algorithm to identify the devices which have the most potential for improvement:

- CPExpert computes the average device response time for each **type** of device in the configuration, for each RMF measurement interval. The logic computes the average device response by type of device, since better performance would be expected from cached devices (for example) than from non-cached devices. This method essentially assesses the performance of each device against the performance of similar devices in the configuration.

- Devices that exceed the average device response time for their device type in any RMF measurement interval are selected as candidates for improvement. The rationale is that improvement efforts should not be directed at devices which provide better than average response. Thus, the candidate set of devices to analyze consists of those which provided worse than average response.

- The I/O rate of each “candidate device is weighted by its response time, **for the entire set of RMF intervals in which the device exceeded the average response**. The result is a measure of the relative performance improvement **potential** of each device which provided worse than average response, from an overall system view. For example, consider two devices in a device type having an average I/O response of 20 milliseconds:

  Device A: I/O rate = 30 I/O operations per second
  Device response = 25 milliseconds
  RMF intervals with above average response = 4
  Seconds per RMF interval = 900
  Weighting factor = 30 * 25 * 4 * 900 = 27,000,000

Device B: I/O rate = 5 I/O operations per second
Device response = 40 milliseconds
RMF intervals with above average response = 5
Seconds per RMF interval = 900
Weighting factor = 5 * 40 = 900,000

In the above example, CPExpert would select Device A as having the most overall potential for improvement, even though its per-I/O device response was not as bad as the device response of Device B.

CPExpert ranks the devices based on the weighting factor computed above. CPExpert then analyzes the devices, starting at the device with the highest weighting factor.

With Rule DAS050, CPExpert lists basic characteristics of the volumes having the most potential for improvement, so that you can appreciate the relative performance improvement potential between volumes on the list. The data presented by Rule DAS050 reflects the average per-second delays **only** during measurement intervals when the device I/O performance was worse than the average for its device type. This information is presented on a system view basis.

The "weighted delays" value is a relative measure of the performance improvement potential of the volume. The absolute values in the column are not particularly meaningful. Rather, the values should be compared to each other to assess the relative performance impact of each volume.

It is possible that a volume may have a significant improvement potential in a particular measurement interval, but not be the volume with the most overall potential for improvement. This situation can arise because the analysis is directed toward the volumes with the **most overall** performance improvement potential. If you suspect that this is the case with a particular device, you can “select” that device for analysis, using the select process described in Section 3 of this document.

The following example illustrates the output from Rule DAS050:

```
        DASD PERFORMANCE CONSTRAINTS ON 22OCT2001 - SYSTEM: SYZ0
                Analyzed by Version 12.1.0 of CPExpert(tm)
          (C) Copyright 2002, Computer Management Sciences, Inc.
                               PAGE: 65


RULE DAS050:  PERFORMANCE CHARACTERISTICS OF SIGNIFICANT VOLUMES

   The following is a list of the most significant volumes showing their
   overall performance characteristics for the period being analyzed.
   The "average per second delays" represent the averages ONLY during
   measurement intervals when the device I/O performance was worse than
   the average for this device type.  The "weighted delays" value is a
   measure of the overall relative performance impact of each volume.

           DEVICE    I/O   ------AVERAGE PER SECOND DELAYS------- WEIGHTED
   VOLSER  NUMBER   RATE    RESP     CONN    DISC    PEND    IOSQ  DELAYS
   SP0006   FE58   114.0 118.340    0.130   0.003   0.785 117.423  236679
   PAGE25   5145     3.8   0.093    0.031   0.054   0.008   0.000     186
   PAGE27   5164     3.7   0.091    0.029   0.053   0.009   0.000     182
   PAGE26   5150     3.8   0.089    0.028   0.052   0.009   0.000     179
   DB007D   FF33     0.4   0.052    0.001   0.017   0.005   0.030     105
   PETNV1   FEE7     2.0   0.045    0.022   0.000   0.022   0.001      90
   DB0053   FEF9     0.3   0.043    0.002   0.001   0.001   0.038      87
   D83PPL   BAD4     1.8   0.037    0.021   0.000   0.016   0.000      75
   CAT011   FD7C     1.2   0.022    0.002   0.009   0.010   0.001      45
   D83PB1   BABA     0.5   0.020    0.006   0.000   0.014   0.001      41
   SS5173   5173     1.7   0.038    0.029   0.007   0.002   0.000      38
```

In this example, it is clear that SP0006 has significant performance improvement potential on SYZ0. The DASD Component would analyze SP0006 as the “worst” device for system SYZ0, to determine what caused the delays. Additionally, if the CPExpert modification to MXG or MICS (described in Section 2) had been installed, the DASD Component would list the applications referencing SP0006. Further, if SMF Type 42 records were available (and the volume contained data sets managed by DFSMS), the DASD Component would produce Rule DAS400 to show access characteristics of the most significant data sets that resided on SP0006.

Suggestion: You should use the information displayed by Rule DAS050 to assess the relative importance of the "worst" performing device compared with the performance improvement potential of the other devices, from the perspective of the “loved one” workload.

Finding: CPExpert identifies the performance characteristics of the volumes in a system that have the most potential for performance improvement, **from the perspective of the “loved one” workload**.

Impact: This finding is used to assess the importance of the "worst" performing device from the perspective of the “loved one” workload, and to determine whether other devices offer significant performance improvement potential.

Logic flow: This is a basic finding. There are no predecessor rules.

Discussion: Rule DAS055 is similar to Rule DAS050, except that Rule DAS055 relates to devices accessed by “loved one” work. Please refer to Rule DAS050 for a discussion of the approach to selecting devices for analysis.

With Rule DAS055, CPExpert lists basic characteristics of the volumes having the most potential for improvement **from the perspective of the “loved one” workload**, so that you can appreciate the relative performance improvement potential between volumes on the list. The data presented by Rule DAS055 reflects the average per-second delays **only** during measurement intervals when the device I/O performance was worse than the average for its device type, **and** for measurement intervals when the device was referenced by the “loved one” workload. This information is presented on a system view basis.

The "weighted delays" value is a relative measure of the performance improvement potential of the volume. The absolute values in the column are not particularly meaningful. Rather, the values should be compared to each other to assess the relative performance impact of each volume.

It is possible that a volume may have a significant improvement potential in a particular measurement interval, but not be the volume with the most overall potential for improvement. This situation can arise because the analysis is directed toward the volumes with the **most overall** performance improvement potential. If you suspect that this is the case with a particular device, you can “select” that device for analysis, using the select process described in Section 3 of this document.

The following example illustrates the output from Rule DAS055 when a “loved one” workload was defined as BATCH:


```
RULE DAS055:  PERFORMANCE CHARACTERISTICS OF SIGNIFICANT VOLUMES

   The following is a list of the most significant volumes accessed by
   BATCH showing their overall performance characteristics for the period
   being analyzed.  The "average per second delays" represent the averages
   ONLY during measurement intervals when the device I/O performance was
   worse than the average for this device type.  The "weighted delays"
   value is a measure of the overall relative performance impact of each
   volume.

           DEVICE   I/O   ------AVERAGE PER SECOND DELAYS-------  WEIGHTED
   VOLSER  NUMBER  RATE    RESP     CONN    DISC    PEND    IOSQ   DELAYS
   WORKP2   2324    32.0   0.320    0.305   0.005   0.008   0.001    8631
   MVSPX1   232E    21.2   0.036    0.029   0.002   0.005   0.000      64
   MVSPL1   2537    73.5   0.132    0.107   0.005   0.016   0.004      59
   MVS902   2137    47.5   0.045    0.031   0.002   0.011   0.001      58
   PRD015   2538    38.5   0.037    0.024   0.004   0.009   0.000      51
   PRD002   252D    16.5   0.016    0.011   0.002   0.003   0.000      50
   PRD004   2536    10.8   0.015    0.011   0.001   0.003   0.000      38
   PRD017   2635    73.8   0.061    0.042   0.003   0.016   0.000      32
   PRD007   2133    10.5   0.011    0.007   0.002   0.002   0.000      32
   WORKP3   2220     7.4   0.072    0.065   0.004   0.002   0.002      30
   PRD005   2038    24.8   0.027    0.018   0.002   0.006   0.000      29
```

In this example, WORKP2 has significant performance improvement potential from the perspective of the BATCH workload[1]. The DASD Component would analyze WORKP2 as the “worst” device for the BATCH workload, to determine what caused the delays. Additionally, if the CPExpert modification to MXG or MICS (described in Section 2) had been installed, the DASD Component would list the applications referencing WORKP2. Further, if SMF Type 42 records were available (and the volume contained data sets managed by DFSMS), the DASD Component would produce Rule DAS400 to show access characteristics of the most significant data sets that resided on WORKP2.

Suggestion: You should use the information displayed by Rule DAS055 to assess the relative importance of the "worst" performing device compared with the performance improvement potential of the other devices, from the perspective of the “loved one” workload.

[1] The BATCH workload was selected simply as an example of Rule DAS055.

Rule DAS100: VOLUME WITH WORST OVERALL PERFORMANCE

Finding: The identified volume had the worst overall performance during the entire measurement period. RULE DAS100 is quite similar to RULE DAS200; RULE DAS100 applies to **all** DASD devices, while RULE DAS200 applies only to DASD devices accessed by critical (or "loved one") workload.

Impact: The impact of this finding will depend upon the importance of the volume to overall system performance. If this is a critical volume, then this finding will have a HIGH impact. However, if the volume is accessed by low priority workloads, then this finding will have a LOW IMPACT or MEDIUM IMPACT.

Address spaces are retained in storage so long as they have uncompleted I/O operations. While the address spaces are in storage, they occupy page frames and may delay other address spaces. Additionally, the SRB time required to service I/O operations executes at a higher dispatching priority than a TCB, regardless of the dispatching priority of the TCB. Thus, there may be an overall system impact even though the volume may be accessed only by low priority workloads.

Logic flow: This is a basic rule finding; there are no predecessor rules.

Discussion: CPExpert determines the average device response time, by device type, for each measurement interval. A "device type" for this purpose is any unique device type (e.g., IBM-3380 or IBM-3390), with the device type modified to reflect whether the device is cached, has DLS or DLSE capabilities, or is a paging device.

The purpose of determining the average device response time, by device type, is the underlying principle that there is little point in analyzing a particular device if its response time is better than average. Rather, the most improvement potential resides with devices whose response time is worse than average.

CPExpert selects a device in each measurement interval for further analysis if the device response time exceeds the average for its device type.

CPExpert consolidates the channel path information (contained in the SMF Type 78 records) with the information contained in SMF Type 74 records. This information, along with the IOCP macro information, allows CPExpert to build a model of the I/O configuration. This model includes all channel path, controller, and device utilization and queuing information. CPExpert then applies queuing formulae to the model to compute delays that occurred at significant parts of the model. The results from the model are associated with essential information describing the device response characteristics.

CPExpert constructs a frequency distribution of all devices whose response is worse than average, weighted by the number of I/O operations executed by the device. This yields a weighted measure of the potential performance improvement that might be achieved for each device. This frequency distribution is sorted descendingly, to yield an ordered list of the devices with the most improvement potential.

CPExpert selects the top devices from the ordered list of devices with the most improvement potential. The results from the model described above is extracted for each of these devices. Detailed information regarding the "worst" device is extracted and reported for each measurement interval. (RULE DAS190 reports summary results for the remaining top devices.)

Exhibit DAS100-1 provides a sample output resulting from the analysis. The VOLSER and device number of the "worst" performing device are identified in the narrative. Information is provided about the overall average I/O rate and the device utilization for the entire measurement period being analyzed.

As shown in Exhibit DAS100-1, CPExpert provides a summary for each measurement interval, showing the average I/O rate for the interval, and the **average delay time per second** during the interval (the total is shown as **I/O RESP** in Exhibit DAS100-1). The average delay time per second effectively reflects the percent of each second (shown in milliseconds) in which the associated delay occurred. The sum of the CONN, DISC, and PEND time is the average device busy time during an average second.

CPExpert analyzes each measurement interval and provides an indication of the major problem in the interval. The problems detected by CPExpert may be (1) excessive seeking, (2) missed RPS reconnect, (3) excessive connect, (4) missed cache hits, or (5) IOS queue. CPExpert determines the major cause of device response delays (simply dividing each potential delay by the total device response time). Additionally, the IOSQ time is divided by the total I/O service time.

The result from these calculations is evaluated to determine whether any area significantly predominates the I/O response time. If so, the respective area is listed as the major problem with the volume during the interval being analyzed. There are certain implications related to this analysis in each area of potential problem:

- **Excessive seeking**. CPExpert computes an average seek delay time by subtracting the average latency and the missed RPS reconnect time from the total disconnect time. The remaining time is seek time for the volume. Excessive seeking is considered to be a major problem if the average seek time accounts for a significant portion of the device service time.

- **Missed RPS reconnect**. CPExpert creates a model of the I/O configuration, and assesses the probability that a missed RPS reconnect will occur (a M/M/C queuing model is used to estimate this probability). Missed RPS reconnect is determined to be a major problem if the number of missed RPS reconnects (multiplied by the device rotation time) accounts for a significant portion of the device service time.

- **Excessive connect**. Excessive connect time is considered to be a major problem if the device connect time accounts for a significant portion of the device service time.

- **IOS queue**. IOS queue time is considered to be a major problem if the IOS queue time accounts for a significant portion of the overall I/O response time. Note that IOS queue time is evaluated against the overall I/O response time. This is done because other problems (e.g., seeking) may cause excessive IOS queue time. The IOS queue time is unlikely to be resolved unless the underlying problem is addressed. Thus, IOS queue time must pass a more stringent test to be considered the major problem.

- **Missed cache hits**. CPExpert estimates the delay due to missed cache hits for cached devices. This estimation is done in the absence of RMF Cache Reporter data, and must be viewed with some caution. If the DISC time is less than the average latency, CPExpert attributes the DISC time to missed cache hits.

  If the DISC time is greater than the average latency, CPExpert attributes the average latency to missed cache hits (the remainder of the DISC time is attributed to missed RPS reconnect and seek). If the estimated time attributed to missed cache hits accounts for a majority of the I/O delay time, CPExpert concludes that the major problem with the volume is cache miss operations.

  The rationale for this approach is that I/O operations to cached devices should not normally experience large average disconnect times. If they do experience large average disconnect times, either (1) many records are not found in the cache or (2) a large number of records require seek operations. If a cached volume experiences large average disconnect times, either the volume is a bad cache candidate or else other volumes cause the cache to be ineffective for the volume.

The device may have no consistent problem. The problems may be concentrated in only a few measurement intervals and these few measurement intervals may dominate the performance characteristics of the volume. The worst of these intervals will be analyzed separately.

Additionally, it is possible that the device has no major problem for any interval. This condition most likely would occur if the guidance provided to CPExpert is too restrictive. For example, if a large number of volumes are excluded from analysis (using the EXCLUDE option), the remaining volumes may have no particular problem. However, the logic is designed to select a "worst" volume, regardless of whether that volume actually has problems. CPExpert tests for this condition and provides appropriate information.

Suggestion: There are no suggestions directly associated with this rule. Subsequent rules will analyze the device problems and attempt to determine the cause of poor performance.

```
RULE DAS100:  VOLUME WITH WORST OVERALL PERFORMANCE

  VOLSER RSA002 (device 0194) had the worst overall performance during
  the entire measurement period (13:59, 23JUL1991 to 16:59, 23JUL1991).
  This pack had an overall average of 17.2 I/O operations per second,
  was busy processing I/O for an average of 47% of the time, and
  had I/O operations queued for an average of 20% of the time.  The
  following summarizes significant performance characteristics of
  VOLSER RSA002:

                          I/O  I/0   --AVERAGE PER SECOND DELAYS--   MAJOR
  MEASUREMENT INTERVAL   RATE  RESP  CONN    DISC    PEND     IOSQ   PROBLEM
  13:59-14:14,23JUL1991  13.7 0.480  0.085   0.223   0.050    0.122  QUEUING
  14:14-14:29,23JUL1991  16.5 0.581  0.102   0.270   0.049    0.160  QUEUING
  14:29-14:44,23JUL1991  16.0 0.552  0.098   0.268   0.059    0.127  SEEKING
  14:44-14:59,23JUL1991  16.5 0.572  0.101   0.289   0.054    0.128  SEEKING
  14:59-15:14,23JUL1991  15.8 0.500  0.099   0.256   0.049    0.096  SEEKING
  15:14-15:29,23JUL1991  16.9 0.665  0.096   0.288   0.073    0.208  QUEUING
  15:29-15:44,23JUL1991  19.0 0.825  0.104   0.347   0.101    0.273  QUEUING
  15:44-15:59,23JUL1991  18.5 0.737  0.103   0.350   0.053    0.231  QUEUING
  15:59-16:14,23JUL1991  16.8 0.688  0.096   0.298   0.061    0.233  QUEUING
  16:14-16:29,23JUL1991  19.9 0.793  0.108   0.354   0.058    0.273  QUEUING
  16:29-16:44,23JUL1991  19.6 0.874  0.102   0.356   0.105    0.311  QUEUING
  16:44-16:59,23JUL1991  17.5 0.669  0.098   0.320   0.061    0.190  QUEUING
```

VOLUME WITH WORST OVERALL PERFORMANCE

EXHIBIT DAS100-1

Rule DAS102: VOLUME WITH NEXT WORST OVERALL PERFORMANCE

Finding: The identified volume had the next worst overall performance during the entire measurement period.

Impact: The impact of this finding will depend upon the importance of the volume to overall system performance. If this is a critical volume, then this finding will have a HIGH impact. However, if the volume is accessed by low priority workloads, then this finding will have a LOW IMPACT or MEDIUM IMPACT.

Address spaces are retained in storage so long as they have uncompleted I/O operations. While the address spaces are in storage, they occupy page frames and may delay other address spaces. Additionally, the SRB time required to service I/O operations executes at a higher dispatching priority than a TCB, regardless of the dispatching priority of the TCB. Thus, there may be an overall system impact even though the volume may be accessed only by low priority workloads.

Logic flow: This is a basic rule finding; there are no predecessor rules.

Discussion: Rule DAS100 identified the volume that had the worst overall performance during the entire measurement period. Rule DAS102 is produced for each successive "worst performing" device selected from an ordered list (Rule DAS190 shows the ordered list of devices).

The number of devices analyzed by Rule DAS102 (and successive rules resulting from the analysis of each device) is controlled by the **ANALYZE** guidance variable (see Section 3: Specifying Guidance Variables).

Please refer to Rule DAS100 for a discussion of the information presented with Rule DAS102.

Suggestion: There are no suggestions directly associated with this rule. Subsequent rules will analyze the device problems and attempt to determine the cause of poor performance.

Rule DAS105: VOLUME PERFORMANCE WAS NOT CONSISTENTLY POOR

Finding: The performance of the volume having the "worst" overall performance was not consistently poor in any single area.

Impact: This finding has little impact, other than to indicate that there may be no problem with the device.

Logic flow: The following rules cause this rule to be invoked:
DAS100: Volume with worst overall performance

Discussion: CPExpert checked seek time, missed RPS reconnect time, PEND time, CONN time, missed cache hits, and IOSQ time. None of these areas **consistently** accounted for a majority of the delay for the volume selected as the volume with the "worst" overall performance.

Suggestion: This condition most likely would occur when the period(s) of poor performance were extremely poor or if the guidance provided to CPExpert is too restrictive.

- Devices sometimes have periods of extremely poor performance combined with a large number of I/O operations. This combination may result in the device being selected as the one with the most potential for performance improvement, even though the device did not have consistently poor performance. The most likely cause of such a situation is a particular application or combination of applications executing in the interval with poor performance.

- The guidance provided to CPExpert may be too restrictive. For example, if a large number of volumes are excluded from analysis (using the EXCLUDE option), the remaining volumes may have no particular problem. However, the logic is designed to select a "worst" volume, regardless of whether that volume actually has problems.

Rule DAS110: SEEKING WAS THE MAJOR CAUSE OF RESPONSE DELAY

Finding: Seeking was the major cause of the I/O response delay with the device.

Impact: This finding can have a MEDIUM IMPACT or HIGH IMPACT, depending upon the importance of the applications referencing the device and on the amount of seeking being done.

Logic flow: The following rule causes this rule to be invoked:
DAS100: Volume with the worst overall performance

Discussion: As described in Chapter 1 of Section 5 of this User Manual, DISC time is composed of seeking, latency, and missed RPS reconnect. The discussion associated with RULE DAS100 describes how CPExpert creates a model of the I/O configuration based upon the IOCP macros and RMF data.

CPExpert applies queuing formulae to the model to estimate the amount of delay attributed to missed RPS reconnect (these delays are a function of the probability of a device finding all paths busy when the device tries to reconnect to the channel path).

The estimated missed RPS reconnect time is subtracted from the DISC time reported by RMF. Additionally, the average latency for the device type is subtracted from the DISC time. The resulting time is assumed to be the seek time. (Note the below discussion about why this assumption might not be correct.)

CPExpert performs the above analysis for each measurement interval reflected in the data. RULE DAS110 is produced if seeking was the major problem for a majority of the measurement intervals.

There are potential problems with this approach, although the approach is generally used throughout the computer industry as a way of estimating missed RPS delays and of estimating seeking.

- The queueing formulae assume exponential interarrival times, exponential service distributions, and an infinite population (the M/M/c formula - Erlang's C formula - is used for the calculations). These assumptions may not be correct if, for example, the I/O activity is a function of a single application.

In his class "MVS I/O Configuration Management", Dr. Jeffery Buzen provides a Dump/Restore application as an excellent example of an application which does not follow standard queuing assumptions.

- The device may be cached, and it may be impossible to apportion the DISC time residual after subtracting missed RPS reconnect time. This time may represent a few missed cache read operations with long seek distances, or may represent a relatively large number missed cache read operations with little seeking but the standard latency for the device.

Thus, the seeking analysis can only show potential problems, rather be considered a definitive indication. However, it is usually a fairly accurate indication of the problem. If high average seeking is reported, you can be fairly certain that high seeking did occur. This is particularly true if the problem is reported throughout the measurement intervals. The "uncertainty" tends to be related to relatively low seeking or seeking reported for cached devices.

RULE DAS110 reports the overall average number of milliseconds out of each second in which the device was positioning the arm. Additionally, RULE DAS110 summarizes key information about the period of worst performance, if seeking was the major cause of delay during this period.

Suggestion: The seeks can be minimized by (1) rearranging files within the pack, (2) moving files from the pack to another actuator, (3) changing the application file accessing characteristics, or (4) possibly restricting the applications allowed to access the pack.

Rule DAS111: SEEKING WAS PROBABLY CAUSED BY INDEPENDENT APPLICATIONS

Finding: CPExpert determined that seeking was the major cause of delay in DASD response for the device. More than half of the I/O queuing to the device was explained using a queuing model. Consequently, CPExpert believes that the seeking probably was caused by independent applications.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device. If the finding is correct, and the independent applications are referencing different files, then the corrective actions could result in significant performance improvements.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS110: Seeking was the major cause of response delay

Discussion: CPExpert uses a M/M/1 queuing model to calculate an estimated queue time for each measurement interval being analyzed. The underlying assumptions of the model are exponential interarrival times, exponential service distributions, and an infinite population. If device activity occurs in this way, the queuing model can predict the expected queuing delays.

If the queuing delay as measured by RMF is significantly different from the estimated delay from the model, it would be clear that the activity did not occur in a random fashion, and most likely the cause of the difference would be that the interarrival times are not randomly distributed.

However, the I/O delay to this pack was fairly well explained by the queuing model. More than half of the I/O queuing delay was explained by the model, for a majority of the measurement intervals. This indicates that the activity was mostly random, as would be expected from independent applications accessing files on the pack.

Suggestion: The most improvement for this pack would likely result from (1) separating the files to different packs, (2) rearranging the files within the pack, (3) tuning the file structure (for example, compressing a shared partitioned data set (PDS), (4) scheduling the applications to avoid contention, or (5) examining the applications doing most of the I/O.

You can identify the files with the most activity in one of several ways, depending upon the volume and the application involved. The method you select should depend upon the availability of information and tools in your environment.

- Discuss the file activity with applications personnel or systems personnel (depending upon the volume). If knowledgeable personnel are available, this is often the quickest and easiest way to determine how to solve the problems. Once a file access problem has been brought to their attention, applications or systems personnel often can easily decide upon a good solution.

- Use an exit available in MXG or in MICS to select Type 30(DD) information just for the volume. The Type 30(DD) information is rarely retained in a performance data base because it is so voluminous. However, you easily can code an exit in MXG or MICS to select the Type 30(DD) information for a specific volume. The amount of data selected would not be too large in most cases.

  You can then write a SAS program to list the DD names used by applications, weighted by the number of accesses. For example, you could code the following to analyze data extracted during MXG update processing:

```
PROC FREQ DATA=pdblib.file;
    WHERE DEVNR = addressX;
    TABLES DDNAME/OUT=TEMP NOPRINT;
    WEIGHT EXCPS;
PROC SORT  DATA=TEMP;
    BY DESCENDING COUNT;
PROC PRINT;
RUN;
```

  The device address is displayed by RULE DAS100, so the "address" in the above coding would be replaced with the actual address displayed by that rule. If you have a MXG performance data base, the address is retained in hexadecimal format, so you would suffix an X to the address.

  If you have a MICS performance data base, the variable names must be changed appropriately (for example, DEVNR would be changed to DEVADDR). Additionally, the address is retained in MICS as a character representation of the value, so you do not need to suffix an X to the address. Rather, you must enclose the address in quotes.

  The result from the above code would be a list of all DDNAMEs referencing the device being analyzed, weighted by the number of

EXCPs to the device, and ordered descendingly with the DD statements for the most active files listed first. You often cannot be certain that the most referenced files are causing problems. However, in most cases, you will find that only a few files (often only 2 or 3) account for over 90% of the accesses. These files generally will be the ones causing arm contention problems.

You can then write a SAS program to list the applications (jobs) referencing the device, weighted by the number of accesses. For example, you could code the following to analyze data extracted during MXG update processing:

```
PROC FREQ DATA=pdblib.file;
    WHERE DEVNR = addressX;
    TABLES JOB/OUT=TEMP NOPRINT;
    WEIGHT EXCPS;
PROC SORT  DATA=TEMP;
    BY DESCENDING COUNT;
PROC PRINT;
RUN;
```

The result from the above code would be a list of all jobs referencing the device being analyzed, weighted by the number of EXCPs to the device, and ordered descendingly with the jobs with the most active files on the device listed first. These jobs generally will be the ones causing arm contention problems. (If you have a MICS performance data base, you would change the code as described earlier.)

- Use a commercially-available DASD activity monitor to isolate the files accessed on the problem volume. If such a monitor is available, this would be a direct way to determine the problems.

Rule DAS112: SEEKING WAS PROBABLY CAUSED BY A SINGLE APPLICATION

Finding: CPExpert determined that seeking was the major cause of delay in DASD response for the device. Less than half of the I/O queuing to the device was explained using a queuing model. Consequently, CPExpert believes that the seeking probably was caused by a single application, rather than by independent applications.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device. If the finding is correct, then actions directed to the specific application could result in significant performance improvements.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS110: Seeking was the major cause of response delay

Discussion: CPExpert uses a M/M/1 queuing model to calculate an estimated queue time for each measurement interval being analyzed. The underlying assumptions of the model are exponential interarrival times, exponential service distributions, and an infinite population. If device activity occurs in this way, the queuing model can predict the expected queuing delays.

If the queuing delay as measured by RMF is significantly different from the estimated delay from the model, it would be clear that the activity did not occur in a random fashion, and most likely the cause of the difference would be that the interarrival times are not randomly distributed.

This was the case with the volume being analyzed. Consequently, CPExpert concludes that the activity was caused by a single application accessing the device. Note that there could be more than one application accessing the device, but if their access patterns were not random with respect to each other, then this conclusion would still be valid.

Suggestion: The most improvement for this pack would likely result from (1) separating the files to different packs, (2) rearranging the files within the pack, (3) tuning the file structure (for example, compressing a shared partitioned data set (PDS), or (4) examining the application doing most of the I/O.

You can identify the files with the most activity in one of several ways, depending upon the volume and the application involved. The method you select should depend upon the availability of information and tools in your environment.

- Discuss the file activity with applications personnel or systems personnel (depending upon the volume). If knowledgeable personnel are available, this is often the quickest and easiest way to determine how to solve the problems. Once a file access problem has been brought to their attention, applications or systems personnel often can easily decide upon a good solution.

- Use an exit available in MXG or in MICS to select Type 30(DD) information just for the volume. The Type 30(DD) information is rarely retained in a performance data base because it is so voluminous. However, you easily can code an exit in MXG or MICS to select the Type 30(DD) information for a specific volume. The amount of data selected would not be too large in most cases.

  You can then write a SAS program to list the DD names used by applications, weighted by the number of accesses. For example, you could code the following to analyze data extracted during MXG update processing:

```
PROC FREQ DATA=pdblib.file;
   WHERE DEVNR = addressX;
   TABLES DDNAME/OUT=TEMP NOPRINT;
 WEIGHT EXCPS;
PROC SORT DATA=TEMP;
   BY DESCENDING COUNT;
PROC PRINT;
RUN;
```

  The device address is displayed by RULE DAS100, so the "address" in the above coding would be replaced with the actual address. If you have a MXG performance data base, the address is retained in hexadecimal format, so you should suffix an X to the address.

  If you have a MICS performance data base, the address is retained as a character representation of the value, so you do not need to suffix an X to the address. Rather, you must enclose the address in quotes.

  The result from the above code would be a list of all DDNAMEs referencing the device being analyzed, weighted by the number of EXCPs to the device, and ordered descendingly with the DD statements for the most active files listed first. You often cannot be certain that the most referenced files are causing problems. However, in most cases, you will find that only a few files (often only 2 or 3) account for over 90% of the accesses. These files generally will be the ones causing arm contention problems.

- Use a commercially-available DASD activity monitor to isolate the files accessed on the problem volume. If such a monitor is available, this would be a direct way to determine the problems.

Rule DAS113: WORST SEEKING WAS PROBABLY CAUSED BY INDEPENDENT APPLICATIONS

Finding: CPExpert determined that seeking was the major cause of delay in DASD response for the device during the measurement interval with the worst I/O response. More than half of the I/O queuing to the device was explained using a queuing model. Consequently, CPExpert believes that the seeking during this interval probably was caused by independent applications.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device. If the finding is correct, and the independent applications are referencing different files, then the corrective actions could result in significant performance improvements.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS110: Seeking was the major cause of response delay

Discussion: CPExpert uses a M/M/1 queuing model to calculate an estimated queue time for the measurement interval with the worst I/O response for the device being analyzed. The underlying assumptions of the model are exponential interarrival times, exponential service distributions, and an infinite population. If device activity occurs in this way, the queuing model can predict the expected queuing delays.

If the queuing delay as measured by RMF is significantly different from the estimated delay from the model, it would be clear that the activity did not occur in a random fashion, and most likely the cause of the difference would be that the interarrival times are not randomly distributed.

However, the I/O delay to this pack was fairly well explained by the queuing model. More than half of the I/O queuing delay was explained by the model, for a majority of the measurement intervals. This indicates that the activity was mostly random, as would be expected from independent applications accessing files on the pack.

Suggestion: The most improvement for this pack would likely result from (1) separating the files to different packs, (2) rearranging the files within the pack, (3) tuning the file structure (for example, compressing a shared partitioned data set (PDS), (4) scheduling the applications to avoid contention, or (5) examining the applications doing most of the I/O.

The actions to be taken when this rule is produced are the same associated with RULE DAS111. Please refer to that rule for further suggestions.

Rule DAS114: WORST SEEKING WAS PROBABLY CAUSED BY A SINGLE APPLICATION

Finding: CPExpert determined that seeking was the major cause of delay in DASD response for the device, during the measurement interval with the worst I/O response. Less than half of the I/O queuing to the device was explained using a queuing model. Consequently, CPExpert believes that the seeking probably was caused by a single application, rather than by independent applications.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device. If the finding is correct, then actions directed to the specific application could result in significant performance improvements.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS110: Seeking was the major cause of response delay

Discussion: CPExpert uses a M/M/1 queuing model to calculate an estimated queue time for the measurement interval with the worst I/O response for the device being analyzed. The underlying assumptions of the model are exponential interarrival times, exponential service distributions, and an infinite population. If device activity occurs in this way, the queuing model can predict the expected queuing delays.

If the queuing delay as measured by RMF is significantly different from the estimated delay from the model, it would be clear that the activity did not occur in a random fashion, and most likely the cause of the difference would be that the interarrival times are not randomly distributed.

This was the case with the volume being analyzed. Consequently, CPExpert concludes that the activity was caused by a single application accessing the device. Note that there could be more than one application accessing the device, but if their access patterns were not random with respect to each other, then this conclusion would still be valid.

Suggestion: The most improvement for this pack would likely result from (1) separating the files to different packs, (2) rearranging the files within the pack, (3) tuning the file structure (for example, compressing a shared partitioned data set (PDS), or (4) examining the application doing most of the I/O.

The actions to be taken when this rule is produced are the same associated with RULE DAS112. Please refer to that rule for further suggestions.

Rule DAS115: SEEKING WAS CAUSE OF I/O DELAY ON PAGE PACK

Finding: CPExpert determined that seeking was the major cause of delay in DASD response for the device with the worst performance, and the device was a paging volume.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device. High seeking on paging volumes potentially can have a significant impact on overall system performance if overall page delay is too high.

Logic flow: The following rules cause this rule to be invoked:
DAS100: Volume with the worst overall performance
DAS110: Seeking was the major cause of response delay

Discussion: CPExpert has determined that seeking was the major cause of I/O response delays on the volume with the worst performance. CPExpert noticed that the device was a page pack.

Suggestion: If this finding resulted from an analysis of your entire DASD environment, then the finding is quite significant.

However, if you had requested an analysis of only the paging volumes (using the "select volumes for analysis" option in DASGUIDE), then the finding may not be significant. Please keep in mind that the logic of the DASD Component will always identify a "worst" device.

You should use the MVS Component or the TSO Component to assess whether paging delays are a major cause of performance degradation.

- If paging delays are not a major cause of performance degradation, then this finding probably should be ignored.

- If paging delays are a major cause of performance degradation, then you should consider adding another page pack to the configuration for this system.

Rule DAS120: MAJOR CAUSE OF I/O DELAY WAS MISSED RPS RECONNECT

Finding: CPExpert has determined that missed Rotational Position Sensing (RPS) reconnects was a major cause of delay in DASD response for the device.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance

Discussion: CPExpert computes the average channel path busy for paths to the device. SMF Type 78 information is used to acquire the channel path utilization, and the IOCP macro information is used to determine the channel paths to the device.

If the device is capable of DLS mode, two channel paths can be concurrently busy to the string. If the device is capable of DLSE mode, four channel paths can be concurrently busy to the string.

CPExpert use an M/M/C queuing model (Erlang's C formula) to compute the probability that all paths to the device were busy when the device attempted to reconnect. The M/M/C queuing model is adjusted depending upon the number of concurrent paths[1]. The average number of missed RPS reconnect attempts is given by the formula:

$$N = \frac{U}{1 - U}$$

where U is the probability that a reconnect attempt will find all paths busy

The average time spent attempting to reconnect is simply the rotational time of the device multiplied by the average number of missed RPS reconnect attempts. This yields an estimate of the average I/O delay caused by missed RPS reconnect attemts.

[1] Please refer to *Probability, Statistics, and Queuing Theory* by Arnold O. Allen for a description of the M/M/C queuing model.

Rule DAS120 is produced if the average time spent attempting to reconnect to the channel path accounts for a significant percent of the device response time.

The I/O delay caused by missed RPS reconnect may be under-estimated if the device is shared between systems. This is because the controller may be busy to a different system when the device attempts to reconnect. There is no information in RMF records to show the time the controller was busy to another system. RMF data from all systems sharing the controller must be consolidated and analyzed in order to obtain this information. The DASD Component does not perform this analysis at present. However, the logic necessary to determine this situation will be added in future updates to the product.

Suggestion: Missed RPS reconnect time is caused by too much activity on the channel paths to the device. This problem can be corrected by:

- Removing I/O data transfer from the path(s). When devices are seeking or searching for a sector, the channel path is not busy. Therefore, high channel activity is primarily due to transferring data (controller protocol accounts for some path busy, but this time generally is quite small). Removing I/O data transfer from the path(s) can be accomplished by:

  - Moving data sets from the paths. This often is the easiest solution, since data sets with high data transfers can be relocated to other volumes on different channel paths.

  - The volumes responsible for high data transfer time could be moved to other strings on different channel paths.

- Rescheduling workloads to minimize the contention. For example, some batch jobs may be performing heavy I/O to volumes on the string. These jobs may be rescheduled to a time when their I/O would not cause problems.

- Adding paths to the device. If the device is not capable of DLS or DLSE mode, this will require upgrading the device (e.g., from IBM-3380 Model A04 to a more recent model). If the device is capable of DLS mode (that is, it can dynamically reconnect to either of two paths), you may consider upgrading the device to one capable of DLSE mode (that is, it can dynamically reconnect to any of four paths). This action may require that you upgrade your controller from an IBM-3880 storage controller to an IBM-3990 storage controller. Since upgrades are relatively expensive, you should first assess the feasibility of moving active data sets or volumes from the path.

- Adding channel paths, acquiring additional controllers, and moving some volumes to the new controllers. This would be a more expensive option, and may not be feasible (depending upon the processor and I/O configuration).

Rule DAS121: VOLUMES CONTRIBUTING TO RPS DELAY

Finding: CPExpert identifies the volumes contributing to missed RPS reconnect delays.

Impact: This information can be useful when deciding on a course of action to correct the missed RPS reconnect problems.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS120: Missed RPS reconnect was major cause of I/O delay

Discussion: If Rule DAS120 is produced, CPExpert examines the SMF Type 74 information and IOCP macro information to select all devices sharing paths with the device experiencing missed RPS reconnect delays. Devices contributing 5% or more to the utilization of the channel paths are listed, ordered descendingly by their contribution to path utilization.

A device contributes to path utilization mostly based upon the connect time of the device. Consequently, CPExpert displays the devices having the largest per-second connect times on the path.

Suggestion: You should consider separating the volumes contributing to missed RPS delays from those on the volume experiencing the significant missed RPS delays.

- You can move data sets from the volumes contributing the most to path utilization.

- You can move volumes to a different string.

- You might reschedule workload to minimize the path utilization at critical times.

Rule DAS123: NON-DASD DEVICES CONTRIBUTED TO RPS DELAY

Finding: CPExpert has determined that non-DASD I/O devices were attached to a channel path of the volume experiencing missed RPS reconnect delays. These non-DASD I/O devices were busy a significant percent of the time and contributed to the RPS delay.

Impact: This finding can have a HIGH IMPACT on the performance of the device experience the missed RPS reconnect delays.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS120: Missed RPS reconnect was major cause of I/O delay

Discussion: CPExpert determines whether any non-DASD I/O devices (e.g., tapes drives, etc.) share channel paths with DASD. If missed RPS reconnect delays were a major cause of I/O delay, CPExpert undertakes an analysis of the non-DASD I/O devices sharing channel paths. CPExpert examines the SMF Type 74 information to determine whether these non-DASD devices had a significant connect time to the path.

CPExpert uses a M/M/c queuing model to estimate the amount of missed RPS reconnect delay caused by the path utilization of the non-DASD devices.

Rule DAS123 is produced if the queuing model estimates that path utilization of the non-DASD devices causes more than 10% of the missed RPS reconnect delay.

Suggestion: CPExpert suggests that you eliminate or minimize the impact of the non-DASD I/O devices on the DASD performance by:

- Consider rescheduling the workload accessing the non-DASD I/O devices to a period when the data transfer would not cause DASD problems.

- Remove the non-DASD I/O devices from the channel paths serving the DASD devices. This may mean that you must acquire additional channel paths.

- If neither of the above options are feasible, consider placing only low-utilization (and non-critical) DASD on the paths shared with the non-DASD I/O devices.

Rule DAS125: APPLICATIONS CONTRIBUTING TO RPS DELAY

Finding: CPExpert identifies the applications contributing to missed RPS reconnect delays to the volume with worst performance.

Impact: This information can be useful when deciding on a course of action to correct the missed RPS reconnect problems. This rule applies only if "expanded" analysis is performed (that is, only if the modification is made to MXG or MICS so that volume information can be associated with workloads).

Logic flow: The following rules cause this rule to be invoked:
DAS100: Volume with the worst overall performance
DAS120: Missed RPS reconnect was major cause of I/O delay

Discussion: If Rule DAS120 is produced, CPExpert examines the SMF Type 30(Interval) information to select all applications which reference the volume. An application contributes to path utilization (and thus, contributes to missed RPS reconnect) mostly based upon the connect time of I/O operations to the device.

CPExpert lists the applications accessing the volume with the worst performance, during the entire measurement period being analyzed. The list is ordered descendingly by the average percent use of paths, so you can identify the applications which probably had the most impact on the volume.

As described in Section 5, the SMF Type 30(Interval) information is not synchronized with the SMF Type 70(series) information. Consequently, the identification of applications may not correctly identify the applications with the most I/O operations to the volume.

However, if the analysis produces the same results after analyzing more than one day's measurement data, you can be more comfortable that the applications are correctly identified.

Suggestion: In many cases, once the applications are identified, either applications personnel or systems personnel will verify whether the particular applications are responsible for the majority of the I/O operations to the particular volume.

You should become comfortable that the applications presented are responsible for the utilization of the paths during the period when missed RPS reconnect was a major cause of device delay. You can then take action to (1) reschedule the applications to minimize contention, (2) examine the files referenced by the applications, or (3) change the way in which the applications access their files.

Finding: CPExpert has determined that excessive PEND time was a major cause of delay in DASD response for the device.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance

Discussion: PEND time is the time from the issuance of the SSCH instruction until the device is selected by the control unit. This time is caused by queuing for the path (wait for channel, wait for director port, wait for control unit, wait for device, or wait for "other" reasons).

CPExpert computes the average per-second PEND delay time, for each of the causes listed above. Rule DAS130 is produced if the average PEND time accounted for a significant percent of the device response time.

Suggestion:Large PEND times usually involve the following situations:

- **Shared devices**. If the device is shared with another system, PEND time may indicate contention with the other system. Large PEND times in shared-device environments usually involve situations very similar to those described under IOSQ time:

  - Multiple data sets may be active on the volume. This situation is the most common and easiest to solve. The data sets can be redistributed among different volumes, to eliminate the queuing at the channel level (reflected as PEND time) for the single volume.
    If some of the data sets are not required to be shared, then the Data Base Administrator has complete flexibility to move these data sets (subject, of course, to the performance implications of the target devices). These data sets should be moved to a non-shared device.

    If the data sets are required to be shared, then they must be relocated to shared devices.

  - Multiple applications or users may be using the same data set on the volume. Depending upon the data set characteristics, duplicate copies of the data set may be placed on different volumes.

This would solve the PEND problems cause by contending systems. If this option is feasible, the data sets could be placed on non-shared devices, likely resulting in even more performance improvement.

- Multiple application systems may be using the volume experiencing high PEND times. In this case, perhaps application redesign or scheduling can solve the problem.

Additionally, large PEND times for shared devices could be caused by RESERVE from the other system. The applications issuing the RESERVE should be examined to determine whether the RESERVE is required. If the RESERVE **is** required, the above situations should reviewed to determine whether improvements can be achieved.

- **Non-shared devices**. Large PEND times for devices that are not shared may mean that there are insufficient paths available to the device. Too much I/O may be directed to many devices on the path, control unit, or head-of-string. The data sets can be redistributed among different volumes on different paths, control units, or heads-of-string. This will reduce the hardware-level queuing. Alternatively, the entire volume may be moved to a different (less busy) head-of-string or path.

  If redistributing the data sets or moving the volume is not feasible, then the device should have more paths. Depending upon the existing configuration, this may involve re-configuring existing channel paths, or acquiring additional hardware.

- **Devices attached to cached controllers**. Large PEND times for devices attached to cached controllers may imply a high percent of read miss operations, or non-volatile storage (NVS) writes for IBM-3990-3 devices. Fairchild [1] lists four ways in which staging in caching controllers can cause hidden device busy (with the device busy potentially reflected in high PEND time):

  - The normal (random) caching algorithm stages all records to the end of the track after a requested record is read.

  - The normal (random) caching algorithm stages all records from the beginning of the track to the requested record if a front-end miss occurs.

[1] Fairchild, Bill, "The Anatomy of an I/O Request", *Conference Proceedings*, CMG'90, the Computer Measurement Group, Chicago, IL.

    - The sequential caching algorithm stages all records to the end of the track after the requested record is read, and stages in all of the next track. IBM-3990 (Model 3) controllers stages in all of the next three tracks.

    - Most writes to extended function IBM-3990 (Model 3) go into NVS with a subsequent destaging required.

- **Dual Copy Initialize**. Large PEND times for IBM-3390 devices may be caused by dual copy initialize. In this case, the dual copy initialize should be turned off.

Rule DAS131: PEND TIME DELAY WAS CAUSED BY PATHS TO DEVICE

Finding: CPExpert has determined that the excessive PEND time was caused by utilization of the paths to the device.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS130: Major cause of I/O delay was PEND time

Discussion: PEND time is the time from the issuance of the SSCH instruction until the device is selected by the control unit. This time is caused by queuing for the path (wait for channel, wait for control unit or wait for head-of-string), and can be caused by other systems sharing the device (wait for device).

Large PEND times for devices that are not shared may mean that there are insufficient paths available to the device. Too much I/O may be directed to many devices on the path, control unit, or head-of-string. CPExpert computes the path utilization for paths to the device. Rule DAS131 is produced if the path utilization for paths to the device is sufficiently large to cause PEND delays.

Suggestion: High PEND caused by path utilization might be solved by redistributed data sets among different volumes on different paths, control units, or heads-of-string. This will reduce the hardware-level queuing. Alternatively, the entire volume may be moved to a different (less busy) head-of-string or path.

If redistributing the data sets or moving the volume is not feasible, then the device should have more paths. Depending upon the existing configuration, this may involve re-configuring existing channel paths, acquiring additional hardware, or upgrading the controller and device so that DLSE (four concurrent paths) can be installed.

Rule DAS132: LARGE PEND TIME MAY BE CAUSED BY OTHER SYSTEM

Finding: A significant amount of the PEND time delay may have been caused by activity from another system sharing the volume, since the PEND time was not explained by high path activity. (Note that other activity not detectable by CPExpert could have caused high PEND time, so the cause may not be activity by another system.)

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS130: Major cause of I/O delay was PEND time

Discussion: PEND time is the time from the issuance of the SSCH instruction until the device is selected by the control unit. This time is caused by queuing for the path (wait for channel, wait for control unit or wait for head-of-string), and can be caused by other systems sharing the device (wait for device).

If the large PEND time was not explained by path utilization, CPExpert determines whether the IOCP macros indicated that the device was shared with another system. If the device is shared with another system, large PEND time may indicate contention with the other system.

Suggestion: Large PEND times in shared-device environments usually involve the following situations:

- Multiple data sets may be active on the volume. This situation is the most common and easiest to solve. The data sets can be redistributed among different volumes, to eliminate the queuing at the channel level (reflected as PEND time) for the single volume.

  If some of the data sets are not required to be shared, then the Data Base Administrator has complete flexibility to move these data sets (subject, of course, to the performance implications of the target devices). These data sets should be moved to a non-shared device.

  If the data sets are required to be shared, then they must be relocated to shared devices.

- Multiple applications or users may be using the same data set on the volume. Depending upon the data set characteristics, duplicate copies of the data set may be placed on different volumes. This would solve the PEND problems cause by contending systems. If this option is feasible, the data sets could be placed on non-shared devices, likely resulting in even more performance improvement.

- Multiple application systems may be using the volume experiencing high PEND times. In this case, perhaps application redesign or scheduling can solve the problem.

- The other system may be issuing a RESERVE command for the device. The applications causing the RESERVE to be issued should be examined to determine whether the RESERVE is required. If the RESERVE **is** required, the above situations should reviewed to determine whether improvements can be achieved.

Rule DAS133: LARGE PEND TIME MAY BE CAUSED BY CONTROLLER ACTIVITY

Finding: A significant amount of the PEND time delay may have been caused by controller activity. This controller activity probably was caused by staging/destaging of cached controllers.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS130: Major cause of I/O delay was PEND time

Discussion: PEND time is the time from the issuance of the SSCH instruction until the device is selected by the control unit. This time is caused by queuing for the path (wait for channel, wait for control unit or wait for head-of-string), and can be caused by other systems sharing the device (wait for device).

Large PEND times for devices attached to cached controllers may imply a high percent of read miss operations, or non-volatile storage (NVS) writes for IBM-3990-3 devices.

Staging in caching controllers can cause hidden device busy (with the device busy potentially reflected in high PEND time):

- The normal (random) caching algorithm stages all records to the end of the track after a requested record is read.

- The normal (random) caching algorithm stages all records from the beginning of the track to the requested record if a front-end miss occurs.

- The sequential caching algorithm stages all records to the end of the track after the requested record is read, and stages in all of the next track. IBM-3990 (Model 3) controllers stage in all of the next three tracks.

- Most writes to extended function IBM-3990 (Model 3) go into NVS with a subsequent destaging required.

Suggestion: Large PEND times caused by staging or destaging of cached controllers implies a low hit rate for data in the cache. There are several alternatives you should consider:

- Determine whether workload scheduling would increase the cache hit rate. Potentially, contending workloads access data sets in a pattern requiring that data be frequently staged into the cache (for read operations) or destaged from the cache (for write operations).

- Examine the cache hit rate of the volumes attached to the controller and potentially remove selected data sets or volumes from the cache, or move them to other cached devices.

  If you have the Cache RMF Reporter (IBM Product 5798-DQD) or comparable products, you will be able to obtain measurement information about the cache utilization and hit rates. The DASD Component of CPExpert does not analyze the Cache RMF Reporter data at present. However, this analysis will be done in a future update to the Component.

- Large PEND times for IBM-3390 devices may be caused by dual copy initialize. In this case, the dual copy initialize should be turned off.

Rule DAS140: CONNECT TIME WAS A MAJOR CAUSE OF I/O DELAY

Finding: Connect time was a major cause of the I/O delay with the volume.

Impact: This finding may have a LOW IMPACT or MEDIUM IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:
DAS100: Volume with the worst overall performance

Discussion: Connect time is the time in which the device is actually connected to the path. This time includes the data transfer time, but also includes protocol exchange (or "hand shaking") between the various components at several stages of the I/O operation.

The data transfer time obviously is a function of the amount of data being transferred. This simply is the number of bytes transferred divided by the transfer speed (for example, if 4096 bytes were transferred from an IBM-3380 with a transfer speed of 3,000,000 bytes per second, the 4096 bytes would require 4096/3,000,000 seconds; or about 1.36 milliseconds).

Large connect times generally are caused by the following situations:

- A large average block size. This situation may be highly desireable for sequential data sets, but would be undesirable for randomly accessed data.

- Long multi-track searches. For example, the catalog must be searched for cataloged files, the Volume Table of Contents (VTOC) must searched to find a requested file, a directory must be searched for partitioned data sets, etc.. These searches will result in long connect times for the volume involved.

- Program loading from system packs.

Rule DAS150: QUEUING IN IOS WAS A MAJOR CAUSE OF I/O DELAY

Finding: Queuing in the I/O Supervisor (IOSQ) was a major cause of the I/O delay with the volume.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:
DAS100: Volume with the worst overall performance

Discussion: IOSQ time is the time from the issuance of a STARTIO macro until the Start SubChannel (SSCH) instruction is issued. After the STARTIO macro is issued, the software determines whether the device is busy. If the device is not busy with this system, the SSCH instruction is issued. However, if the device is busy with this system, the I/O request is queued. Thus, IOSQ time always means that the device is unable to handle additional requests from this system.

Some small IOSQ time is often unavoidable. However, large IOSQ time imply a situation that should be examined. Large IOSQ times result from (1) too many I/O operations directed to the device or (2) lengthy device response times (perhaps caused by high seeking, high RPS delays, or high PEND time).

Suggestion: Large IOSQ times usually involve the following situations:

- Multiple data sets may be active on the volume. This situation is the most common and easiest to solve. The data sets can be redistributed among different volumes, to eliminate the queuing for the single volume.

- Multiple users may be using the same data set on the volume. Depending upon the data set characteristics, duplicate copies of the data set placed on different volumes may solve the IOSQ problems.

- Multiple application systems may be using the volume experiencing high IOSQ times. In this case, perhaps application redesign or scheduling can solve the problem.

- A particular application (or system function) may be executing I/O to the device faster than the device can respond.

- The overall device response time (PEND, DISC, and CONN) times may be large, such that the device is unable to provide quick response to the I/O requests. This situation will be revealed by large values in the PEND, DISC, or CONN measures.

Rule DAS151: QUEUING WAS PROBABLY CAUSED BY INDEPENDENT APPLICATIONS

Finding: More than half of the I/O queuing to the device was explained using a queuing model. Consequently, CPExpert believes that the queuing probably was caused by independent applications.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device. If the finding is correct, and the independent applications are referencing different files, then the corrective actions could result in significant performance improvements.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS150: Queuing was the major cause of response delay

Discussion: CPExpert uses a M/M/1 queuing model to calculate an estimated queue time for each measurement interval being analyzed. The underlying assumptions of the model are exponential interarrival times, exponential service distributions, and an infinite population. If device activity occurs in this way, the queuing model can predict the expected queuing delays.

If the queuing delay as measured by RMF is significantly different from the estimated delay from the model, it would be clear that the activity did not occur in a random fashion, and most likely the cause of the difference would be that the interarrival times are not randomly distributed.

However, the I/O delay to this pack was fairly well explained by the queuing model. More than half of the I/O queuing delay was explained by the model, for a majority of the measurement intervals. This indicates that the activity was mostly random, as would be expected from independent applications accessing files on the pack.

Suggestion: The most improvement for this pack would likely result from (1) separating the files to different packs, (2) rearranging the files within the pack, (3) tuning the file structure (for example, compressing a shared partitioned data set (PDS), (4) scheduling the applications to avoid contention, or (5) examining the applications doing most of the I/O.

If CPExpert is performing "expanded" analysis, Rule DAS180 will provide information regarding the applications referencing the volume. The results from this rule should be examined to determine which applications are candidates for actions.

You can identify the files with the most activity in one of several ways, depending upon the volume and the application involved. The method you select should depend upon the availability of information and tools in your environment.

- Discuss the file activity with applications personnel or systems personnel (depending upon the volume). If knowledgeable personnel are available, this is often the quickest and easiest way to determine how to solve the problems. Once a file access problem has been brought to their attention, applications or systems personnel often can easily decide upon a good solution.

- Use an exit available in MXG or in MICS to select Type 30(DD) information just for the volume. The Type 30(DD) information is rarely retained in a performance data base because it is so voluminous. However, you easily can code an exit in MXG or MICS to select the Type 30(DD) information for a specific volume. The amount of data selected would not be too large in most cases.

  You can then write a SAS program to list the DD names used by applications, weighted by the number of accesses. For example, you could code the following to analyze data extracted during MXG update processing:

  ```
  PROC FREQ DATA=pdblib.file;
      WHERE DEVNR = addressX;
      TABLES DDNAME/OUT=TEMP NOPRINT;
      WEIGHT EXCPS;
  PROC SORT DATA=TEMP;
      BY DESCENDING COUNT;
  PROC PRINT;
  RUN;
  ```

  The device address is displayed by RULE DAS100, so the "address" in the above coding would be replaced with the actual address displayed by that rule. If you have a MXG performance data base, the address is retained in hexadecimal format, so you would suffix an X to the address. If you have a MICS performance data base, the variable names must be changed appropriately (for example, DEVNR would be changed to DEVADDR). Additionally, the address is retained in MICS as a character representation of the value, so you do not need to suffix an X to the address. Rather, you must enclose the address in quotes.

The result from the above code would be a list of all DDNAMEs referencing the device being analyzed, weighted by the number of EXCPs to the device, and ordered descendingly with the DD statements for the most active files listed first. You often cannot be certain that the most referenced files are causing problems. However, in most cases, you will find that only a few files (often only 2 or 3) account for over 90% of the accesses. These files generally will be the ones causing arm contention problems.

You can then write a SAS program to list the applications (jobs) referencing the device, weighted by the number of accesses. For example, you could code the following to analyze data extracted during MXG update processing:

```
PROC FREQ DATA=pdblib.file;
    WHERE DEVNR = addressX;
    TABLES JOB/OUT=TEMP NOPRINT;
    WEIGHT EXCPS;
PROC SORT DATA=TEMP;
    BY DESCENDING COUNT;
PROC PRINT;
RUN;
```

The result from the above code would be a list of all jobs referencing the device being analyzed, weighted by the number of EXCPs to the device, and ordered descendingly with the jobs with the most active files on the device listed first. These jobs generally will be the ones causing arm contention problems. (If you have a MICS performance data base, you would change the code as described earlier.)

- Use a commercially-available DASD activity monitor to isolate the files accessed on the problem volume. If such a monitor is available, this would be a direct way to determine the problems.

Rule DAS152: QUEUING WAS PROBABLY CAUSED BY A SINGLE APPLICATION

Finding: Less than half of the I/O queuing to the device was explained using a queuing model. Consequently, CPExpert believes that the seeking probably was caused by a single application, rather than by independent applications.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device. If the finding is correct, then actions directed to the specific application could result in significant performance improvements.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS150: Queuing was the major cause of response delay

Discussion: CPExpert uses a M/M/1 queuing model to calculate an estimated queue time for each measurement interval being analyzed. The underlying assumptions of the model are exponential interarrival times, exponential service distributions, and an infinite population. If device activity occurs in this way, the queuing model can predict the expected queuing delays.

If the queuing delay as measured by RMF is significantly different from the estimated delay from the model, it would be clear that the activity did not occur in a random fashion, and most likely the cause of the difference would be that the interarrival times are not randomly distributed.

This was the case with the volume being analyzed. Consequently, CPExpert concludes that the activity was caused by a single application accessing the device. Note that there could be more than one application accessing the device, but if their access patterns were not random with respect to each other, then this conclusion would still be valid.

Suggestion: The most improvement for this pack would likely result from (1) separating the files to different packs, (2) rearranging the files within the pack, (3) tuning the file structure (for example, compressing a shared partitioned data set (PDS), or (4) examining the application doing most of the I/O.

You can identify the files with the most activity in one of several ways, depending upon the volume and the application involved. The method you select should depend upon the availability of information and tools in your environment.

- Discuss the file activity with applications personnel or systems personnel (depending upon the volume). If knowledgeable personnel are available, this is often the quickest and easiest way to determine how to solve the problems. Once a file access problem has been brought to their attention, applications or systems personnel often can easily decide upon a good solution.

- Use an exit available in MXG or in MICS to select Type 30(DD) information just for the volume. The Type 30(DD) information is rarely retained in a performance data base because it is so voluminous. However, you easily can code an exit in MXG or MICS to select the Type 30(DD) information for a specific volume. The amount of data selected would not be too large in most cases.

- If CPExpert is performing "expanded" analysis, Rule DAS180 will provide information regarding the applications referencing the volume. The results from this rule should be examined to determine which applications are candidates for actions.

Rule DAS153: WORST QUEUING WAS PROBABLY CAUSED BY INDEPENDENT APPLICATIONS

Finding: CPExpert determined that seeking was the major cause of delay in DASD response for the device during the measurement interval with the worst I/O response. More than half of the I/O queuing to the device was explained using a queuing model. Consequently, CPExpert believes that the queuing during this interval probably was caused by independent applications.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device. If the finding is correct, and the independent applications are referencing different files, then the corrective actions could result in significant performance improvements.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS150: Seeking was the major cause of response delay

Discussion: CPExpert uses a M/M/1 queuing model to calculate an estimated queue time for the measurement interval with the worst I/O response for the device being analyzed. The underlying assumptions of the model are exponential interarrival times, exponential service distributions, and an infinite population. If device activity occurs in this way, the queuing model can predict the expected queuing delays.

If the queuing delay as measured by RMF is significantly different from the estimated delay from the model, it would be clear that the activity did not occur in a random fashion, and most likely the cause of the difference would be that the interarrival times are not randomly distributed.

However, the I/O delay to this pack was fairly well explained by the queuing model. More than half of the I/O queuing delay was explained by the model, for a majority of the measurement intervals. This indicates that the activity was mostly random, as would be expected from independent applications accessing files on the pack.

Suggestion: The most improvement for this pack would likely result from (1) separating the files to different packs, (2) rearranging the files within the pack, (3) tuning the file structure (for example, compressing a shared partitioned data set (PDS), (4) scheduling the applications to avoid contention, or (5) examining the applications doing most of the I/O.

The actions to be taken when this rule is produced are the same associated with RULE DAS151. Please refer to that rule for further suggestions.

Rule DAS154: WORST QUEUING WAS PROBABLY CAUSED BY A SINGLE APPLICATION

Finding: CPExpert determined that IOSQ was the major cause of delay in DASD response for the device, during the measurement interval with the worst I/O response. Less than half of the I/O queuing to the device was explained using a queuing model. Consequently, CPExpert believes that the queuing probably was caused by a single application, rather than by independent applications.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device. If the finding is correct, then actions directed to the specific application could result in significant performance improvements.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS150: Seeking was the major cause of response delay

Discussion: CPExpert uses a M/M/1 queuing model to calculate an estimated queue time for the measurement interval with the worst I/O response for the device being analyzed. The underlying assumptions of the model are exponential interarrival times, exponential service distributions, and an infinite population. If device activity occurs in this way, the queuing model can predict the expected queuing delays.

If the queuing delay as measured by RMF is significantly different from the estimated delay from the model, it would be clear that the activity did not occur in a random fashion, and most likely the cause of the difference would be that the interarrival times are not randomly distributed.

This was the case with the volume being analyzed. Consequently, CPExpert concludes that the activity was caused by a single application accessing the device. Note that there could be more than one application accessing the device, but if their access patterns were not random with respect to each other, then this conclusion would still be valid.

Suggestion: The most improvement for this pack would likely result from (1) separating the files to different packs, (2) rearranging the files within the pack, (3) tuning the file structure (for example, compressing a shared partitioned data set (PDS), or (4) examining the application doing most of the I/O.

The actions to be taken when this rule is produced are the same associated with RULE DAS152. Please refer to that rule for further suggestions.

Rule DAS160: MAJOR CAUSE OF I/O DELAY WAS MISSED CACHE READ HITS

Finding: CPExpert determined that missed cache read hits was the major cause of delay in DASD response for the device.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device and on the performance of other volumes attached to the cache controller.

Logic flow: The following rules cause this rule to be invoked:
DAS100: Volume with the worst overall performance

Discussion: Cache controllers offer the potential of **significant** reduction in I/O response times. Cache controllers operate by maintaining some amount of data from the volumes being cached in high-speed memory. The following summarizes the cache controller operation and performance considerations:

- When an I/O read operation is executed for data from a cached volume, the controller determines whether the required data resides in the cache.

  - If the data **does** reside in the cache, then the data is immediately provided. This situation is titled a "read hit" operation. A read hit eliminates the requirement for interaction with the volume (seek operation is potentially eliminated and rotational position sensing is eliminated in all cases). The cost of the read operation is the small amount of controller time involved in determining whether the data resides in the cache and the data transfer time. The I/O read operation savings resulting from a read hit can be 90% or more of the normal read I/O time!

  - If the data does **not** reside in the cache, then the data is retrieved from the volume in the normal read process (i.e., potentially a seek, a set sector, a search, and data transfer). This situation is titled a "read miss" operation. The cost of this read operation is the normal read I/O cost, plus the controller time determining whether the data resided in cache.

    Additionally, upon a read miss, the cache controller will transfer all data from one or more tracks from the volume into cache (the number of tracks transferred depends upon the controller model).

This process is called "staging" of the data, and is done with the idea that if data is read from a file, it is likely that the next sequential records in the file will be read soon. The controller stages these records into cache so that they will be in the cache when a subsequent read operation occurs.

Depending upon the controller model, the controller may be busy during the staging of data. Regardless of the controller model, the device is busy during the staging operation.

- When an I/O write operation is executed for data from a cached volume, the controller determines whether the required data resides in the cache just is it does for a read I/O operation.

  - If the data **does** reside in the cache, then the data is transferred to the cache and the data is transferred from the cache to the volume. This situation is titled a "write hit" operation. A write hit requires transfer of data to the cache and transfer of data to the device in the normal write process (i.e., potentially a seek, a set sector, a search, and data transfer). The cost of this write operation is the channel time to transfer the data to cache, plus the controller time determining whether the data resided in cache.

  - If the data does **not** reside in the cache, then the data is written to the volume in the normal write process (i.e., potentially a seek, a set sector, a search, and data transfer). This situation is titled a "write miss" operation. The cost of this write operation is the normal write I/O cost, plus the controller time determining whether the data resided in cache.

From the above brief discussion, it is clear that the savings from cache controller occurs when the read I/O operations find the required data in the cache. Performance is significantly improved in this situation. Performance suffers because of cache in all other situations. Consequently, it highly desireable to cache volumes which mostly have read hit situations.

In the absence of direct measurements of cache hit/miss operations[1], CPExpert makes some conclusions about the suitability of cache hit/miss based upon the average DISC time of I/O operations to each volume. (Recall that the DISC time is mostly time spent seeking, normal rotational latency, and missed RPS connect.)

- If cache is being used effectively, most I/O operations will be satisfied

[1]Cache hit and miss information is available, for example, from IBM's RMF Cache Reporter.

by read hits and the average DISC time will be quite small. This is because most operations will require no access to the actual device, but will be satisfied from cache. Consequently, if the average DISC time is small, it is clear that the cache is effective.

- If cache is not being used effectively, relatively few I/O operations will be satisfied by read hits, and a relatively large number of I/O operations will require access to the device. These I/O operations will require the normal latency (an average of one-half the rotational time of the device), plus may require seeking to position the arm. Consequently, if the average DISC time is relatively large for cached volumes, one interpretation is that cache is not being effective for the volume.

  CPExpert estimates the effect of missed cache hits. If the DISC time is less than the average latency, CPExpert attributes the DISC time to missed cache hits. If the DISC time is greater than the average latency, CPExpert attributes the average latency to missed cache hits (the remainder of the DISC time is attributed to missed RPS reconnect and seek). If the estimated time attributed to missed cache hits accounts for a majority of the I/O delay time, CPExpert concludes that the major problem with the volume is cache miss operations.

  This analysis approach works well in many situations. However, there is a potential problem with this approach since it is based upon averages. Suppose that a relatively few I/O operations resulted in seeks over a large number of cylinders. The seek time (and thus the DISC time) would be large for these relatively few I/O operations, and these large times would distort the average DISC for all operations. It is possible that most I/O operations to the volume were satisfied by read hits, but the average DISC time would be large because of the long seek operations.

Suggestion: There are several alternatives you should consider:

- If the applications and data contained on the volume are not critical, you should consider moving the volume to a non-cached controller, since it does not appear that the volume significantly benefits from being cached.

- Alternatively, there may be non-critical volumes on the cache controller with data access patterns which cause the cache to be full with their data. These volumes may prevent the cache from adequately responding to the I/O requirements of the volume identified by this rule. You may wish to identify these volumes and move them to a non-cached controller.

- Alternatively, applications accessing the volume may be generating seeks which require movement of the arm over a large number of cylinders. If this is the case, then the actions under Rule DAS110 should be reviewed.

- Alternatively, you may wish to consider acquiring another cached controller if all volumes currently being cached are critical and the I/O requests to the identified volume are **not** being delayed because of long seeks.

Rule DAS170: THERE DID NOT APPEAR TO BE A PROBLEM WITH VOLUME PERFORMANCE

Finding: The performance of the volume having the "worst" overall performance was not poor in any area evaluated by CPExpert.

Impact: This finding has little impact, other than to indicate that there may be no problem with the device.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with worst overall performance

Discussion: CPExpert checked seek time, missed RPS reconnect time, PEND time, CONN time, missed cache hits, and IOSQ time. None of these areas accounted for the delay for the volume selected as the volume with the "worst" overall performance.

Suggestion: This condition most likely would occur when the guidance provided to CPExpert is too restrictive. For example, if a large number of volumes are excluded from analysis (using the EXCLUDE option), the remaining volumes may have no particular problem. However, the logic is designed to select a "worst" volume, regardless of whether that volume actually has problems.

Rule DAS180: APPLICATIONS ACCESSING VOLUME WITH WORST PERFORMANCE

Finding: CPExpert identifies the applications accessing the volume with the worst performance.

Impact: This finding is used to select applications for potential rescheduling, file movement, or design changes. This rule is applicable only if CPExpert is performing "expanded" analysis of the data.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with worst overall performance

Discussion: CPExpert lists the applications accessing the volume with the worst performance during the entire measurement period being analyzed. The list is ordered descendingly by I/O operations, so you can identify the applications which probably had the most impact on the volume.

As described in Section 5, the SMF Type 30(Interval) information is not synchronized with the SMF Type 70(series) information. Consequently, the identification of applications may not correctly identify the applications with the most I/O operations to the volume.

However, if the analysis produces the same results after analyzing more than one day's measurement data, you can be more comfortable that the applications are correctly identified.

Suggestion: In many cases, once the applications are identified, either applications personnel or systems personnel will verify whether the particular applications are responsible for the majority of the I/O operations to the particular volume.

Once you are comfortable that the applications are responsible for the majority of the I/O operations to the "worst" performing device, you can take action to (1) reschedule the applications to minimize contention, (2) examine the files referenced by the applications, or (3) change the way in which the applications access their files.

Rule DAS185: APPLICATIONS ACCESSING VOLUME DURING MEASUREMENT INTERVAL WITH WORST PERFORMANCE

Finding: CPExpert identifies the applications accessing the volume during the the measurement interval with the worst performance.

Impact: This finding is used to select applications for potential rescheduling, file movement, or design changes. This rule is applicable only if CPExpert is performing "expanded" analysis of the data.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with worst overall performance

Discussion: CPExpert lists the applications accessing the volume with the worst performance during the measurement period with the worst performance being analyzed. The list is ordered descendingly by I/O operations, so you can identify the applications which probably had the most impact on the volume.

As described in Section 5, the SMF Type 30(Interval) information is not synchronized with the SMF Type 70(series) information. Consequently, the identification of applications may not correctly identify the applications with the most I/O operations to the volume.

However, if the analysis produces the same results after analyzing more than one day's measurement data, you can be more comfortable that the applications are correctly identified.

Suggestion: In many cases, once the applications are identified, either applications personnel or systems personnel will verify whether the particular applications are responsible for the majority of the I/O operations to the particular volume.

Once you are comfortable that the applications are responsible for the majority of the I/O operations to the "worst" performing device, you can take action to (1) reschedule the applications to minimize contention, (2) examine the files referenced by the applications, or (3) change the way in which the applications access their files.

Rule DAS190: PERFORMANCE CHARACTERISTICS OF SIGNIFICANT VOLUMES

Finding: CPExpert identifies the performance characteristics of the most significant volumes (including the volume selected as the "worst" performing volume).

Impact: This finding is used to assess the importance of the "worst" performing device and to determine whether other devices offer significant performance improvement potential.

Logic flow: This is a basic finding. There are no predecessor rules.

Discussion: CPExpert lists basic characteristics of the volumes having the most potential for improvement. The list includes the volume selected as the "worst" performing device, so that you can appreciate the relative performance improvement potential between the "worst" volume and other volumes on the list.

Volumes are selected as having improvement potential only if their response time exceeds the average for their device type. This "screening" is done for each measurement interval being analyzed. The data presented by Rule DAS190 reflects the average per-second delays only during measurement intervals when the device I/O performance was worse than the average for its device type.

The "weighted delays" value is a relative measure of the performance improvement potential of the volume. The absolute values in the column are not particularly meaningful. Rather, the values should be compared to each other to assess the relative performance impact of each volume.

For example, the "worst" volume might have a "weighted delays" value of 1000. If the "next worst" device had a "weighted delays" value of 950, you may wish to direct CPExpert to examine the next worst device in more detail (this would be accomplished by using DASGUIDE to "exclude" the worst volume).

On the other hand, suppose the "next worst" device had a "weighted delays" value of only 400. You probably would not wish to examine the next worst in more detail, since little performance benefit would be gained from tuning actions directed to the next worst volume.

It is possible that a volume may have a significant improvement potential in a particular measurement interval, but not be the volume with the most overall potential for improvement. This situation can arise because the analysis is directed toward the volumes with the **most overall** performance improvement potential.

Suggestion: You should use the information displayed by Rule DAS190 to assess the relative importance of the "worst" performing device compared with the performance improvement potential of the other devices.

Rule DAS200: VOLUME PROVIDING WORST OVERALL PERFORMANCE TO CRITICAL WORKLOAD

Finding: The identified volume provided the worst overall performance to the "loved one" workload during the entire measurement period. RULE DAS200 is quite similar to RULE DAS100; RULE DAS200 applies only to DASD devices accessed by critical (or "loved one") workload, while RULE DAS100 applies to **all** DASD devices.

Impact: This finding will have a HIGH IMPACT on the performance of the "loved one" workload.

Logic flow: This is a basic rule finding; there are no predecessor rules.

Discussion: CPExpert determines the average device response time, by device type, for each measurement interval. A "device type" for this purpose is any unique device type (e.g., IBM-3380 or IBM-3390), with the device type modified to reflect whether the device is cached, has DLS or DLSE capabilities, or is a paging device.

The purpose of determining the average device response time, by device type, is the underlying principle that there is little point in analyzing a particular device if its response time is better than average. Rather, the most improvement potential resides with devices whose response time is worse than average.

CPExpert selects a device in each measurement interval for further analysis if the device response time exceeds the average for its device type **and** the device was referenced by the "loved one" workload.

CPExpert consolidates the channel path information (contained in the SMF Type 78 records) with the information contained in SMF Type 74 records. This information, along with the IOCP macro information, allows CPExpert to build a model of the I/O configuration. This model includes all channel path, controller, and device utilization and queuing information.

CPExpert then applies queuing formulae to the model to compute delays that occurred at significant parts of the model. The results from the model are associated with essential information describing the device response characteristics.

CPExpert constructs a frequency distribution of all devices whose response is worse than average, weighted by the number of I/O operations executed by the device. This yields a weighted measure of the potential performance improvement that might be achieved for each device. This frequency distribution is sorted descendingly, to yield an ordered list of the devices with the most improvement potential.

CPExpert selects the top devices from the ordered list of devices referenced by the "loved one" workload. These represent the devices with the most improvement potential with respect to the "loved one" workload. The result from the model described above is extracted for each of these devices. Detailed information regarding the "worst" device is extracted and reported for each measurement interval. (RULE DAS290 reports summary results for the remaining top devices.)

Please refer to Section 5 for a discussion regarding some of the limitations of the technique of associating workload to device performance. In brief summary, the Type 30 information is not synchronized with the Type 70(series) information. Consequently, the analysis is approximate, rather than precise.

Exhibit DAS200-1 provides a sample output resulting from the analysis. The VOLSER and device number of the "worst" performing device are identified in the narrative. Information is provided about the overall average I/O rate and the device utilization for the entire measurement period being analyzed.

As shown in Exhibit DAS200-1, CPExpert provides a summary for each measurement interval, showing the average I/O rate for the interval, and the **average delay time per second** during the interval (the total is shown as **I/O RESP** in Exhibit DAS100-1). The average delay time per second effectively reflects the percent of each second (shown in milliseconds) in which the associated delay occurred. The sum of the CONN, DISC, and PEND time is the average device busy time during an average second.

CPExpert analyzes each measurement interval and provides an indication of the major problem in the interval. The problems detected by CPExpert may be (1) excessive seeking, (2) missed RPS reconnect, (3) excessive connect, (4) missed cache hits, or (5) IOS queue. CPExpert determines the major cause of device response delays (simply dividing each potential delay by the total device response time). Additionally, the IOSQ time is divided by the total I/O service time.

The result from these calculations is evaluated to determine whether any area significantly predominates the I/O response time. If so, the respective area is listed as the major problem with the volume during the interval being analyzed. There are certain implications related to this analysis in each area of potential problem:

- **Excessive seeking**. CPExpert computes an average seek delay time by subtracting the average latency and the missed RPS reconnect time from the total disconnect time. The remaining time is seek time for the volume. Excessive seeking is considered to be a major problem if the average seek time accounts for a significant portion of the device service time.

- **Missed RPS reconnect**. CPExpert creates a model of the I/O configuration, and assesses the probability that a missed RPS reconnect will occur (a M/M/C queuing model is used to estimate this probability). Missed RPS reconnect is determined to be a major problem if the number of missed RPS reconnects (multiplied by the device rotation time) accounts for a significant portion of the device service time.

- **Excessive connect**. Excessive connect time is considered to be a major problem if the device connect time accounts for a significant portion of the device service time.

- **IOS queue**. IOS queue time is considered to be a major problem if the IOS queue time accounts for a significant portion of the overall I/O response time. Note that IOS queue time is evaluated against the overall I/O response time. This is done because other problems (e.g., seeking) may cause excessive IOS queue time. The IOS queue time is unlikely to be resolved unless the underlying problem is addressed. Thus, IOS queue time must pass a more stringent test to be considered the major problem.

- **Missed cache hits**. CPExpert estimates the delay due to missed cache hits for cached devices. This estimation is done in the absence of RMF Cache Reporter data, and must be viewed with some caution. If the DISC time is less than the average latency, CPExpert attributes the DISC time to missed cache hits. If the DISC time is greater than the average latency, CPExpert attributes the average latency to missed cache hits (the remainder of the DISC time is attributed to missed RPS reconnect and seek).

  If the estimated time attributed to missed cache hits accounts for a majority of the I/O delay time, CPExpert concludes that the major problem with the volume is cache miss operations.

  The rationale for this approach is that I/O operations to cached devices should not normally experience large average disconnect times. If they do experience large average disconnect times, either (1) many records are not found in the cache or (2) a large number of records require seek operations. If a cached volume experiences large average disconnect times, either the volume is a bad cache candidate or else other volumes cause the cache to be ineffective for the volume.

The device may have no consistent problem. The problems may be concentrated in only a few measurement intervals and these few measurement intervals may dominate the performance characteristics of the volume. The worst of these intervals will be analyzed separately.

Additionally, it is possible that the device has no major problem for any interval. This condition most likely would occur if the guidance provided to CPExpert is too restrictive. For example, if a large number of volumes are excluded from analysis (using the EXCLUDE option), the remaining volumes may have no particular problem. However, the logic is designed to select a "worst" volume, regardless of whether that volume actually has problems. CPExpert tests for this condition and provides appropriate information.

Suggestion: There are no suggestions directly associated with this rule. Subsequent rules will analyze the device problems and attempt to determine the cause of poor performance.

```
RULE DAS200:  VOLUME WITH WORST OVERALL PERFORMANCE

  VOLSER RSA002 (device 0194) provided CICS with the worst overal performance
  during the entire measurement period (13:59, 23JUL1991 to 16:59, 23JUL1991).
  This pack had an overall average of 17.2 I/O operations per second, was busy
  processing I/O for an average of 47% of the time, and had I/O operations
  queued for an average of 20% of the time.  The following summarizes
  significant performance characteristics of VOLSER RSA002:


                          I/O  I/0   --AVERAGE PER SECOND DELAYS--   MAJOR
  MEASUREMENT INTERVAL   RATE  RESP  CONN    DISC    PEND     IOSQ   PROBLEM
  13:59-14:14,23JUL1991  13.7 0.480  0.085   0.223   0.050    0.122  QUEUING
  14:14-14:29,23JUL1991  16.5 0.581  0.102   0.270   0.049    0.160  QUEUING
  14:29-14:44,23JUL1991  16.0 0.552  0.098   0.268   0.059    0.127  SEEKING
  14:44-14:59,23JUL1991  16.5 0.572  0.101   0.289   0.054    0.128  SEEKING
  14:59-15:14,23JUL1991  15.8 0.500  0.099   0.256   0.049    0.096  SEEKING
  15:14-15:29,23JUL1991  16.9 0.665  0.096   0.288   0.073    0.208  QUEUING
  15:29-15:44,23JUL1991  19.0 0.825  0.104   0.347   0.101    0.273  QUEUING
  15:44-15:59,23JUL1991  18.5 0.737  0.103   0.350   0.053    0.231  QUEUING
  15:59-16:14,23JUL1991  16.8 0.688  0.096   0.298   0.061    0.233  QUEUING
  16:14-16:29,23JUL1991  19.9 0.793  0.108   0.354   0.058    0.273  QUEUING
  16:29-16:44,23JUL1991  19.6 0.874  0.102   0.356   0.105    0.311  QUEUING
  16:44-16:59,23JUL1991  17.5 0.669  0.098   0.320   0.061    0.190  QUEUING
```

VOLUME WITH WORST OVERALL PERFORMANCE

EXHIBIT DAS200-1

Rule DAS202: VOLUME PROVIDING NEXT WORST OVERALL PERFORMANCE TO CRITICAL WORKLOAD

Finding: The identified volume provided the next worst overall performance to the "loved one" workload during the entire measurement period.

Impact: This finding will have a HIGH IMPACT on the performance of the "loved one" workload.

Logic flow: This is a basic rule finding; there are no predecessor rules.

Discussion: Rule DAS202 identified the volume that had the worst overall performance during the entire measurement period, from the perspective of the "loved one" workload. Rule DAS202 is produced for each successive "worst performing" device selected from an ordered list (Rule DAS290 shows the ordered list of devices).

The number of devices analyzed by Rule DAS202 (and successive rules resulting from the analysis of each device) is controlled by the **ANALYZE** guidance variable (see Section 3: Specifying Guidance Variables).

Please refer to Rule DAS202 for a discussion of the information presented with Rule DAS202.

Suggestion: There are no suggestions directly associated with this rule. Subsequent rules will analyze the device problems and attempt to determine the cause of poor performance.

Rule DAS205: VOLUME PERFORMANCE WAS NOT CONSISTENTLY POOR

Finding: The performance of the volume having the "worst" overall performance for the critical workload was not consistently poor in any single area.

Impact: This finding has little impact, other than to indicate that there may be no problem with the device.

Logic flow: The following rules cause this rule to be invoked:
DAS200: Volume with worst overall performance

Discussion: CPExpert checked seek time, missed RPS reconnect time, PEND time, CONN time, missed cache hits, and IOSQ time. None of these areas **consistently** accounted for a majority of the delay for the volume selected as the volume with the "worst" overall performance with respect to the critical workload.

Suggestion: This condition most likely would occur when the period(s) of poor performance were extremely poor or if the guidance provided to CPExpert is too restrictive.

- Devices sometimes have periods of extremely poor performance combined with a large number of I/O operations. This combination may result in the device being selected as the one with the most potential for performance improvement, even though the device did not have consistently poor performance. The most likely cause of such a situation is a particular application or combination of applications executing in the interval with poor performance.

- The guidance provided to CPExpert may be too restrictive. For example, if a large number of volumes are excluded from analysis (using the EXCLUDE option), the remaining volumes may have no particular problem. However, the logic is designed to select a "worst" volume, regardless of whether that volume actually has problems.

Rule DAS210: SEEKING WAS THE MAJOR CAUSE OF RESPONSE DELAY

Finding: Seeking was the major cause of the I/O response delay with the device.

Impact: This finding can have a MEDIUM IMPACT or HIGH IMPACT. Since the device is the "worst performing" device for the critical (or "loved one" workload, this finding can have a HIGH IMPACT, depending upon the amount of seeking being done.

Logic flow: The following rule causes this rule to be invoked:

DAS200: Volume with the worst overall performance

Discussion: As described in Chapter 1 of Section 5 of this User Manual, DISC time is composed of seeking, latency, and missed RPS reconnect. The discussion associated with RULE DAS200 describes how CPExpert creates a model of the I/O configuration based upon the IOCP macros and RMF data.

CPExpert applies queuing formulae to the model to estimate the amount of delay attributed to missed RPS reconnect (these delays are a function of the probability of a device finding all paths busy when the device tries to reconnect to the channel path).

The estimated missed RPS reconnect time is subtracted from the DISC time reported by RMF. Additionally, the average latency for the device type is subtracted from the DISC time. The resulting time is assumed to be the seek time. (Note the below discussion about why this assumption might not be correct.)

CPExpert performs the above analysis for each measurement interval reflected in the data. RULE DAS210 is produced if seeking was the major problem for a majority of the measurement intervals.

There are potential problems with this approach, although the approach is generally used throughout the computer industry as a way of estimating missed RPS delays and of estimating seeking.

- The queueing formulae assume exponential interarrival times, exponential service distributions, and an infinite population (the M/M/c formula - Erlang's C formula - is used for the calculations). These assumptions may not be correct if, for example, the I/O activity is a function of a single application.

In his class "MVS I/O Configuration Management", Dr. Jeffery Buzen provides a Dump/Restore application as an excellent example of an application which does not follow standard queuing assumptions.

- The device may be cached, and it may be impossible to apportion the DISC time residual after subtracting missed RPS reconnect time. This time may represent a few missed cache read operations with long seek distances, or may represent a relatively large number missed cache read operations with little seeking but the standard latency for the device.

Thus, the seeking analysis can only show potential problems, rather be considered a definitive indication. However, it is usually a fairly accurate indication of the problem. If high average seeking is reported, you can be fairly certain that high seeking did occur. This is particularly true if the problem is reported throughout the measurement intervals. The "uncertainty" tends to be related to relatively low seeking or seeking reported for cached devices.

RULE DAS210 reports the overall average number of milliseconds out of each second in which the device was positioning the arm **while servicing I/O requests for the "loved one" workload**. It is important to remember that the device may have had different performance characteristics at different measurement intervals when the device was not used by the "loved one" workload.

Additionally, RULE DAS210 summarizes key information about the period of worst performance, if seeking was the major cause of delay during this period.

Suggestion: The seeks can be minimized by (1) rearranging files within the pack, (2) moving files from the pack to another actuator, (3) changing the application file accessing characteristics, or (4) possibly restricting the applications allowed to access the pack.

Rule DAS220: MAJOR CAUSE OF I/O DELAY TO CRITICAL WORKLOAD WAS MISSED RPS RECONNECT

Finding: CPExpert has determined that missed Rotational Position Sensing (RPS) reconnects was a major cause of delay in DASD response for the device.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:

DAS200: Volume with the worst overall performance

Discussion: CPExpert computes the average channel path busy for paths to the device. SMF Type 78 information is used to acquire the channel path utilization, and the IOCP macro information is used to determine the channel paths to the device.

If the device is capable of DLS mode, two channel paths can be concurrently busy to the string. If the device is capable of DLSE mode, four channel paths can be concurrently busy to the string.

CPExpert use an M/M/C queuing model (Erlang's C formula) to compute the probability that all paths to the device were busy when the device attempted to reconnect. The M/M/C queuing model is adjusted depending upon the number of concurrent paths. (Please refer to *Probability, Statistics, and Queuing Theory* by Arnold O. Allen for a description of the M/M/C queuing model.)

The average number of missed RPS reconnect attempts is given by the formula:

$$N = \frac{U}{1 - U}$$

where U is the probability that a reconnect attempt will find all paths busy

The average time spent attempting to reconnect is simply the rotational time of the device multiplied by the average number of missed RPS reconnect attempts. This yields an estimate of the average I/O delay caused by missed RPS reconnect attemts.

Rule DAS220 is produced if the average time spent attempting to reconnect to the channel path accounts for a significant percent of the device response time **to the "loved one" workload**.

The I/O delay caused by missed RPS reconnect may be under-estimated if the device is shared between systems. This is because the controller may be busy to a different system when the device attempts to reconnect. There is no information in RMF records to show the time the controller was busy to another system. RMF data from all systems sharing the controller must be consolidated and analyzed in order to obtain this information. This analysis is performed by the DASD Component only if you have selected the “analysis of shared DASD conflicts” option (see Section 3, Chapter 2.5).

Suggestion: Missed RPS reconnect time is caused by too much activity on the channel paths to the device. This problem can be corrected by:

- Removing I/O data transfer from the path(s). When devices are seeking or searching for a sector, the channel path is not busy. Therefore, high channel activity is primarily due to transferring data (controller protocol accounts for some path busy, but this time generally is quite small). Removing I/O data transfer from the path(s) can be accomplished by:

  - Moving data sets from the paths. This often is the easiest solution, since data sets with high data transfers can be relocated to other volumes on different channel paths.

  - The volumes responsible for high data transfer time could be moved to other strings on different channel paths.

- Rescheduling workloads to minimize the contention. For example, some batch jobs may be performing heavy I/O to volumes on the string. These jobs may be rescheduled to a time when their I/O would not cause problems.

- Adding paths to the device. If the device is not capable of DLS or DLSE mode, this will require upgrading the device (e.g., from IBM-3380 Model A04 to a more recent model). If the device is capable of DLS mode (that is, it can dynamically reconnect to either of two paths), you may consider upgrading the device to one capable of DLSE mode (that is, it can dynamically reconnect to any of four paths). This action may require that you upgrade your controller from an IBM-3880 storage controller to an IBM-3990 storage controller. Since upgrades are relatively expensive, you should first assess the feasibility of moving active data sets or volumes from the path.

- Adding channel paths, acquiring additional controllers, and moving some volumes to the new controllers. This would be a more expensive option, and may not be feasible (depending upon the processor and I/O configuration).

Rule DAS221: VOLUMES CONTRIBUTING TO RPS DELAY

Finding: CPExpert identifies the volumes contributing to missed RPS reconnect delays to the "loved one" workload.

Impact: This information can be useful when deciding on a course of action to correct the missed RPS reconnect problems.

Logic flow: The following rules cause this rule to be invoked:
DAS200: Volume with the worst overall performance
DAS220: Missed RPS reconnect was major cause of I/O delay

Discussion: If Rule DAS220 is produced, CPExpert examines the SMF Type 74 information and IOCP macro information to select all devices sharing paths with the device experiencing missed RPS reconnect delays. Up to 10 of these devices are listed, ordered descendingly by their contribution to path utilization.

A device contributes to path utilization mostly based upon the connect time of the device. Consequently, CPExpert displays the devices having the largest per-second connect times on the path.

Suggestion: You should consider separating the volumes contributing to missed RPS delays from those on the volume experiencing the significant missed RPS delays.

- You can move data sets from the volumes contributing the most to path utilization.

- You can move volumes to a different string.

- You might reschedule workload to minimize the path utilization at critical times.

Rule DAS223: NON-DASD DEVICES CONTRIBUTED TO RPS DELAY

Finding: CPExpert has determined that non-DASD I/O devices were attached to a channel path of the volume experiencing missed RPS reconnect delays. These non-DASD I/O devices were busy a significant percent of the time and contributed to the RPS delay.

Impact: This finding can have a HIGH IMPACT on the performance of the device experience the missed RPS reconnect delays. Since this device is used by the "loved one" workload and is the volume with the worst overall performance, this finding can have a HIGH IMPACT on the performance of the "loved one" workload.

Logic flow: The following rules cause this rule to be invoked

DAS200: Volume with the worst overall performance
DAS220: Missed RPS reconnect was major cause of I/O delay

Discussion: CPExpert determines whether any non-DASD I/O devices (e.g., tapes drives, etc.) share channel paths with DASD. If missed RPS reconnect delays were a major cause of I/O delay, CPExpert undertakes an analysis of the non-DASD I/O devices sharing channel paths. CPExpert examines the SMF Type 74 information to determine whether these non-DASD devices had a significant connect time to the path.

CPExpert uses a M/M/c queuing model to estimate the amount of missed RPS reconnect delay caused by the path utilization of the non-DASD devices.

Rule DAS223 is produced if the queuing model estimates that path utilization of the non-DASD devices causes more than 10% of the missed RPS reconnect delay.

Suggestion: CPExpert suggests that you eliminate or minimize the impact of the non-DASD I/O devices on the DASD performance by considering the following alternatives:

- Reschedule the workload accessing the non-DASD I/O devices to a period when the data transfer would not cause DASD problems.

- Remove the non-DASD I/O devices from the channel paths serving the DASD devices. This may mean that you must acquire additional channel paths.

- If neither of the above options are feasible, consider placing only low-utilization (and non-critical) DASD on the paths shared with the non-DASD I/O devices.

Rule DAS225: APPLICATIONS CONTRIBUTING TO RPS DELAY

Finding: CPExpert identifies the applications (other than the "loved one" workload) contributing to missed RPS reconnect delays.

Impact: This information can be useful when deciding on a course of action to correct the missed RPS reconnect problems.

Logic flow: The following rules cause this rule to be invoked:

DAS200: Volume with the worst overall performance
DAS220: Missed RPS reconnect was major cause of I/O delay

Discussion: If Rule DAS220 is produced, CPExpert examines the SMF Type 30(Interval) information to select all applications (other than the "loved one" workload) which reference the volume. An application contributes to path utilization (and thus, contributes to missed RPS reconnect) mostly based upon the connect time of I/O operations to the device.

CPExpert lists the applications accessing the volume with the worst performance for the "loved one" workload, during the entire measurement period being analyzed. The list is ordered descendingly by the average percent use of paths, so you can identify the applications which probably had the most impact on the volume.

As described in Section 5, the SMF Type 30(Interval) information is not synchronized with the SMF Type 70(series) information. Consequently, the identification of applications may not correctly identify the applications with the most I/O operations to the volume.

However, if the analysis produces the same results after analyzing more than one day's measurement data, you can be more comfortable that the applications are correctly identified.

Suggestion: In many cases, once the applications are identified, either applications personnel or systems personnel will verify whether the particular applications are responsible for the majority of the I/O operations to the particular volume.

You should become comfortable that the applications presented are responsible for the utilization of the paths during the period when missed RPS reconnect delayed the "loved one" workload. You can then take action to (1) reschedule the applications to minimize contention, (2) examine the files referenced by the applications, or (3) change the way in which the applications access their files.

Rule DAS230: MAJOR CAUSE OF I/O DELAY WAS PEND TIME

Finding: CPExpert has determined that excessive PEND time was a major cause of delay in DASD response for the device.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:

DAS200: Volume with the worst overall performance

Discussion: PEND time is the time from the issuance of the SSCH instruction until the device is selected by the control unit. This time is caused by queuing for the path (wait for channel, wait for control unit or wait for head-of-string), and can be caused by other systems sharing the device (wait for device).

CPExpert computes the average per-second PEND delay time. Rule DAS230 is produced if the average PEND time accounted for a significant percent of the device response time.

Suggestion: Please refer to Rule DAS130 for suggestions on how to reduce PEND time.

Rule DAS231: PEND TIME DELAY WAS CAUSED BY PATHS TO DEVICE

Finding: CPExpert has determined that the excessive PEND time was caused by utilization of the paths to the device.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:

DAS200: Volume with the worst overall performance
DAS230: Major cause of I/O delay was PEND time

Discussion: PEND time is the time from the issuance of the SSCH instruction until the device is selected by the control unit. This time is caused by queuing for the path (wait for channel, wait for control unit or wait for head-of-string), and can be caused by other systems sharing the device (wait for device).

Large PEND times for devices that are not shared may mean that there are insufficient paths available to the device. Too much I/O may be directed to many devices on the path, control unit, or head-of-string. CPExpert computes the path utilization for paths to the device. Rule DAS231 is produced if the path utilization for paths to the device is sufficiently large to cause PEND delays.

Suggestion: High PEND caused by path utilization might be solved by redistributed data sets among different volumes on different paths, control units, or heads-of-string. This will reduce the hardware-level queuing. Alternatively, the entire volume may be moved to a different (less busy) head-of-string or path.

If redistributing the data sets or moving the volume is not feasible, then the device should have more paths. Depending upon the existing configuration, this may involve re-configuring existing channel paths, acquiring additional hardware, or upgrading the controller and device so that DLSE (four concurrent paths) can be installed.

Rule DAS232: LARGE PEND TIME MAY BE CAUSED BY OTHER SYSTEM

Finding: A significant amount of the PEND time delay may have been caused by activity from another system sharing the volume, since the PEND time was not explained by high path activity. (Note that other activity not detectable by CPExpert could have caused high PEND time, so the cause may not be activity by another system.)

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:

DAS200: Volume with the worst overall performance
DAS230: Major cause of I/O delay was PEND time

Discussion: PEND time is the time from the issuance of the SSCH instruction until the device is selected by the control unit. This time is caused by queuing for the path (wait for channel, wait for control unit or wait for head-of-string), and can be caused by other systems sharing the device (wait for device).

If the large PEND time was not explained by path utilization, CPExpert determines whether the IOCP macros indicated that the device was shared with another system. If the device is shared with another system, large PEND time may indicate contention with the other system.

Suggestion: Please refer tp Rule DAS131 for suggestions on how to reduce PEND time caused by other systems.

Rule DAS233: LARGE PEND TIME MAY BE CAUSED BY CONTROLLER ACTIVITY

Finding: A significant amount of the PEND time delay may have been caused by controller activity. This controller activity probably was caused by staging/destaging of cached controllers.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:

DAS200: Volume with the worst overall performance
DAS230: Major cause of I/O delay was PEND time

Discussion: PEND time is the time from the issuance of the SSCH instruction until the device is selected by the control unit. This time is caused by queuing for the path (wait for channel, wait for control unit or wait for head-of-string), and can be caused by other systems sharing the device (wait for device).

Please refer to Rule DAS133 for a discussion regarding PEND time caused by controller activity.

Suggestion: Please refer tp Rule DAS133 for suggestions on how to reduce PEND time caused by controller activity.

Rule DAS240: CONNECT TIME WAS A MAJOR CAUSE OF I/O DELAY

Finding: Connect time was a major cause of the I/O delay with the volume.

Impact: This finding may have a LOW IMPACT or MEDIUM IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:
DAS200: Volume with the worst overall performance

Discussion: Connect time is the time in which the device is actually connected to the path. This time includes the data transfer time, but also includes protocol exchange (or "hand shaking") between the various components at several stages of the I/O operation.

The data transfer time obviously is a function of the amount of data being transferred. This simply is the number of bytes transferred divided by the transfer speed (for example, if 4096 bytes were transferred from an IBM-3380 with a transfer speed of 3,000,000 bytes per second, the 4096 bytes would require 4096/3,000,000 seconds; or about 1.36 milliseconds).

Large connect times generally are caused by the following situations:

- A large average block size. This situation may be highly desireable for sequential data sets, but would be undesirable for randomly accessed data.
- Long multi-track searches. For example, the catalog must be searched for cataloged files, the Volume Table of Contents (VTOC) must searched to find a requested file, a directory must be searched for partitioned data sets, etc.. These searches will result in long connect times for the volume involved.
- Program loading from system packs.

Rule DAS250: QUEUING IN IOS WAS A MAJOR CAUSE OF I/O DELAY

Finding: Queuing in the I/O Supervisor (IOSQ) was a major cause of the I/O delay with the volume.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the device.

Logic flow: The following rules cause this rule to be invoked:
DAS200: Volume with the worst overall performance

Discussion: IOSQ time is the time from the issuance of a STARTIO macro until the Start SubChannel (SSCH) instruction is issued. After the STARTIO macro is issued, the software determines whether the device is busy. If the device is not busy with this system, the SSCH instruction is issued. However, if the device is busy with this system, the I/O request is queued. Thus, IOSQ time always means that the device is unable to handle additional requests from this system. (The emphasis on "this system" is explained in the below discussion of PEND time.)

Some small IOSQ time is often unavoidable. However, large IOSQ time imply a situation that should be examined. Large IOSQ times result from (1) too many I/O operations directed to the device or (2) lengthy device response times (perhaps caused by high seeking, high RPS delays, or high PEND time).

Suggestion: Please refer to Rule DAS150 for suggestions on how to reduce high queuing in IOS.

Rule DAS260: MAJOR CAUSE OF I/O DELAY TO CRITICAL APPLICATION WAS MISSED CACHE READ HITS

Finding: CPExpert determined that missed cache read hits was the major cause of delay in DASD response to critical applications for the device.

Impact: This finding may have a MEDIUM IMPACT or HIGH IMPACT on the performance of the critical application, on the device, and on the performance of other volumes attached to the cache controller.

Logic flow: The following rules cause this rule to be invoked:

DAS200: Volume with the worst overall performance

Discussion: Cache controllers offer the potential of **significant** reduction in I/O response times. Cache controllers operate by maintaining some amount of data from the volumes being cached in high-speed memory. Please refer to Rule DAS160 for a discussion regarding the performance of cache controllers.

Suggestion: There are several alternatives you should consider:

- There may be other volumes on the cache controller with data access patterns which cause the cache to be full with their data. These volumes may prevent the cache from adequately responding to the I/O requirements of the volume identified by this rule. You may wish to identify these volumes and move them to a non-cached controller.

- Alternatively, applications accessing the volume may be generating seeks which require movement of the arm over a large number of cylinders. If this is the case, then the actions under Rule DAS110 should be reviewed.

- Alternatively, you may wish to consider acquiring another cached controller if all volumes currently being cached are critical and the I/O requests to the identified volume are **not** being delayed because of long seeks.

Rule DAS270: THERE DID NOT APPEAR TO BE A PROBLEM WITH VOLUME PERFORMANCE

Finding: The performance of the volume having the "worst" overall performance was not poor in any area evaluated by CPExpert.

Impact: This finding has little impact, other than to indicate that there may be no problem with the device.

Logic flow: The following rules cause this rule to be invoked:

DAS200: Volume with worst overall performance

Discussion: CPExpert checked seek time, missed RPS reconnect time, PEND time, CONN time, missed cache hits, and IOSQ time. None of these areas accounted for the delay for the volume selected as the volume with the "worst" overall performance.

Suggestion: This condition most likely would occur when the guidance provided to CPExpert is too restrictive. For example, if a large number of volumes are excluded from analysis (using the EXCLUDE option), the remaining volumes may have no particular problem. However, the logic is designed to select a "worst" volume, regardless of whether that volume actually has problems.

Rule DAS280: APPLICATIONS ACCESSING VOLUME WITH WORST PERFORMANCE

Finding: CPExpert identifies the applications (other than the "loved one" workload accessing the volume with the worst performance.

Impact: This finding is used to select applications for potential rescheduling, file movement, or design changes. This rule is applicable only if CPExpert is performing "expanded" analysis of the data.

Logic flow: The following rules cause this rule to be invoked:

DAS200: Volume with worst overall performance

Discussion: CPExpert lists the applications (other than the "loved one" workload) accessing the volume with the worst performance during the entire measurement period being analyzed. The list is ordered descendingly by I/O operations, so you can identify the applications which probably had the most impact on the volume.

As described in Section 5, the SMF Type 30(Interval) information is not synchronized with the SMF Type 70(series) information. Consequently, the identification of applications may not correctly identify the applications with the most I/O operations to the volume.

However, if the analysis produces the same results after analyzing more than one day's measurement data, you can be more comfortable that the applications are correctly identified.

Suggestion: In many cases, once the applications are identified, either applications personnel or systems personnel will verify whether the particular applications are responsible for the majority of the I/O operations to the particular volume.

Once you are comfortable that the applications are responsible for the majority of the I/O operations to the "worst" performing device, you can take action to (1) reschedule the applications to minimize contention, (2) examine the files referenced by the applications, or (3) change the way in which the applications access their files.

Rule DAS285: APPLICATIONS ACCESSING VOLUME DURING MEASUREMENT INTERVAL WITH WORST PERFORMANCE

Finding: CPExpert identifies the applications (other than the "loved one" workload) accessing the volume during the the measurement interval with the worst performance.

Impact: This finding is used to select applications for potential rescheduling, file movement, or design changes.

Logic flow: The following rules cause this rule to be invoked:
DAS200: Volume with worst overall performance

Discussion: CPExpert lists the applications accessing the volume with the worst performance during the measurement period with the worst performance being analyzed. The list is ordered descendingly by I/O operations, so you can identify the applications which probably had the most impact on the volume.

As described in Section 5, the SMF Type 30(Interval) information is not synchronized with the SMF Type 70(series) information. Consequently, the identification of applications may not correctly identify the applications with the most I/O operations to the volume.

However, if the analysis produces the same results after analyzing more than one day's measurement data, you can be more comfortable that the applications are correctly identified.

Suggestion: In many cases, once the applications are identified, either applications personnel or systems personnel will verify whether the particular applications are responsible for the majority of the I/O operations to the particular volume.

Once you are comfortable that the applications are responsible for the majority of the I/O operations to the "worst" performing device, you can take action to (1) reschedule the applications to minimize contention, (2) examine the files referenced by the applications, or (3) change the way in which the applications access their files.

Rule DAS287: OTHER APPLICATIONS DID NOT REFERENCE THE VOLUME

Finding: CPExpert has determined that only the "loved one" workload referenced the volume during the interval when the device had the worst performance.

Impact: This finding has little impact, other than to indicate that the device performance is a function of the "loved one" workload, rather than performance problems being caused by other workload.

Logic flow: The following rules cause this rule to be invoked:

DAS200: Volume with worst overall performance

Discussion: CPExpert examined the workload accessing the device during the period of "worst" performance. Only the "loved one" workload accessed the device during this measurement interval.

As described in Section 5, the SMF Type 30(Interval) information is not synchronized with the SMF Type 70(series) information. Consequently, the identification of applications may not correctly identify the applications with the most I/O operations to the volume.

However, if the analysis produces the same results after analyzing more than one day's measurement data, you can be more comfortable that other workload did not reference the device.

Suggestion: This finding implies that performance problems for the device are not caused by contending workloads. Performance tuning actions must be directed toward the "loved one" workload.

Rule DAS290: PERFORMANCE CHARACTERISTICS OF SIGNIFICANT VOLUMES

Finding: CPExpert identifies the performance characteristics of the most significant volumes accessed by the "loved one" workload (including the volume selected as the "worst" performing volume).

Impact: This finding is used to assess the importance of the "worst" performing device from the perspective of the "loved one" workload, and to determine whether other devices offer significant performance improvement potential.

Logic flow: This is a basic finding. There are no predecessor rules.

Discussion: CPExpert lists basic characteristics of the volumes having the most potential for improvement. The list includes the volume selected as the "worst" performing device, so that you can appreciate the relative performance improvement potential between the "worst" volume and other volumes on the list.

Volumes are selected as having improvement potential only if their response time exceeds the average for their device type. This "screening" is done for each measurement interval being analyzed. The data presented by Rule DAS190 reflects the average per-second delays only during measurement intervals when the device I/O performance was worse than the average for its device type.

The "weighted delays" value is a relative measure of the performance improvement potential of the volume. The absolute values in the column are not particularly meaningful. Rather, the values should be compared to each other to assess the relative performance impact of each volume.

For example, the "worst" volume might have a "weighted delays" value of 1000. If the "next worst" device had a "weighted delays" value of 950, you may wish to direct CPExpert to examine the next worst device in more detail (this would be accomplished by using DASGUIDE to "exclude" the worst volume).

On the other hand, suppose the "next worst" device had a "weighted delays" value of only 400. You probably would not wish to examine the next worst in more detail, since little performance benefit would be gained from tuning actions directed to the next worst volume.

It is possible that a volume may have a significant improvement potential in a particular measurement interval, but not be the volume with the most overall potential for improvement. This situation can arise because the analysis is directed toward the volumes with the **most overall** performance improvement potential.

Suggestion: You should use the information displayed by Rule DAS190 to assess the relative importance of the "worst" performing device compared with the performance improvement potential of the other devices.

Rule DAS300: Perhaps shared DASD conflicts caused performance problems

Finding: CPExpert believes that accessing conflicts caused by sharing DASD between systems or MVS images may have caused performance problems.

Impact: This finding is used to assess whether sharing DASD between systems or MVS images caused performance problems.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS110: Seeking was major cause of I/O response delay
DAS120: Missed RPS reconnect was major cause of I/O delay
DAS130: Large PEND time was major cause of I/O delay
DAS150: Missed cache read hits was major cause of I/O delay

Discussion: DASD volumes can be shared between systems or between MVS images. Sharing of the DASD volumes might be implemented to allow backup of data, to facilitate recovery or restart, to permit transfer of data from one system to another, etc.

In some situations, sharing DASD volumes has no impact on performance (for example, few I/O operations might be directed to the shared volumes from potentially conflicting systems).

In other situations, sharing DASD volumes can have a significant impact on the performance of the shared volumes, and consequently, on the performance of the applications accessing the shared volumes.

CPExpert can perform an analysis of conflicts between DASD shared between systems or MVS images. The analysis performed by CPExpert is not intended to identify an isolated performance problem. Rather, CPExpert attempts to identify those problems which **continually** cause shared DASD performance problems.

Assume that CPExpert is analyzing DASD performance problems with System A, and some DASD devices are shared with System B. There are three major ways in which the shared DASD can cause performance problems for System A:

- **Case #1**. The arm of the device has been moved by System B when System A attempts to access a cylinder. The DASD I/O operation from System A must move the arm to the desired cylinder, as a SEEK operation.

CPExpert computes the time required to perform SEEKs on System A. If the computed SEEK time is a major cause of performance problems, CPExpert analyzes the data from System B to determine whether System B generates a large number of I/O operations to the device.

- If System B does **not** generate a relatively large number of I/O operations to the device, CPExpert concludes that there is **not** a conflict. There is little doubt about the validity of this conclusion. If System B does not direct much I/O activity to the device, System B clearly cannot cause seek problems for System A.

- If System B **does** generate a relatively large number of I/O operations to the device, CPExpert concludes that there **is** a conflict caused by sharing the device.

  To be absolutely correct, CPExpert should process the configuration definitions for System B, process System B's channel and device information, and compute seek information for System B. CPExpert could then determine whether System B also experienced a high seek rate for the device.

  If both System A and System B experienced a high seek rate, CPExpert could be absolutely sure that there was a shared DASD conflict. This approach would unnecessarily use system resources and would be cumbersome to implement.

  Consequently, CPExpert makes the assumption that I/O operations to the device are random between System A and System B. CPExpert thus can conclude that if System A experiences a high seek rate and System B significantly uses the device (exhibited by a high I/O rate), then System B must also experience a high seek rate. To assume otherwise would require that I/O from System B be coordinated with the I/O from System A, such that System B does not experience seeking similar to System A.

  Since System A experiences a high seek rate and System B experiences a high I/O rate, CPExpert concludes that there is a conflict caused by the shared DASD.

- **Case #2**. The device (or the path to the device) is busy to System B when System A attempts to reference the device. The DASD I/O operation from System A must wait until the device (or path) is free. This wait time is directly reflected as PEND time in the RMF measurement data.

CPExpert determines whether PEND time is a major cause of performance problems for the device (or devices) selected for detailed analysis. If PEND time is a major cause of performance problems, CPExpert analyzes the data from System B to determine whether DASD I/O operations from System B could cause the PEND delays for System A.

CPExpert analyzes the total CONN time and DISC time for the device from other systems sharing the device. CPExpert concludes that the PEND delay is caused by other systems if the total CONN time and DISC time is more than 25% of the PEND delay for the device experienced on System A.

This analysis does not detect PEND delays caused by path delays (e.g., the controller being busy). In order to analyze all path delays, CPExpert would have to process the IOCP macros for all systems and perform a complete analysis of all other I/O configurations. This seems to be an excessive amount of analysis for only marginal gain.

Not detecting PEND delays caused by path conflicts should not be a serious deficiency in the analysis. For most system, the PEND time is not caused by path delays (most systems have sufficient paths for shared DASD that path delays from another system are not serious). The PEND delays would normally be caused by the other system using the device. Consequently, the analysis which is performed should detect most shared DASD problems exhibited by PEND delays.

- **Case #3**. The device is cached, the cached controller is shared between System A and System B, and I/O operations from System B preempt data in the cache which is required by System A.

  This problem is evidenced by relatively large I/O times for System A. Since the data required by System A is not in the cache, the data must be acquired from the device. This potentially results in a seek operation to position the arm, and will require normal latency and data transfer from the device.

  In any of the above situations with cached shared devices, CPExpert analyzes the amount of I/O operations from System A and System B to the cached controller.

- If System B does **not** generate a relatively large number of I/O operations to the device, CPExpert concludes that there is **not** a conflict.

- If System B **does** generate a relatively large number of I/O operations to the device, CPExpert concludes that there **is** a conflict caused by sharing the device. The rationale for these conclusions is the same as was discussed under Case #1 (seek conflicts).

The standard analysis performed by CPExpert may not detect a cache problem under two possible scenarios.

- If only one cached controller is attached to System A, CPExpert may not detect a problem with the device. This is because the logic employed by CPExpert selects devices with the most performance improvement within each **type of device** and then selects the **overall** "worst" devices for detailed analysis.

  CPExpert considers cached devices to be a unique type. If all devices on the cached controller received bad service caused by shared cached problems, CPExpert may not detect a performance problem with the cached devices. This is because all devices in the "device type" could have roughly equal poor service and no device would be **significantly** worse than the other devices in the device type. Consequently, CPExpert might not select any of the cached devices for detailed analysis

  If there are multiple cached controllers, there would be a larger number of "candidate" devices, and the standard analysis performed by CPExpert is more likely to identify any problem caused by shared cache controllers.

- The analysis performed by CPExpert may not detect a cache problem if the cache is being replaced with data from another volume. It is possible that System B could cause data from another volume to be loaded into the controller's cache. This volume could be a volume not be flagged as the "worst" performing volume when analyzing performance from the perspective of System A. From System A's perspective, accesses to the worst volume would simply not find required data in cache.

  Without analyzing the IOCP information for all systems, CPExpert cannot determine which volumes are attached to which controllers. Consequently, CPExpert cannot at present relate (1) poor performance for one volume on System A and (2) I/O operations to a different volume by System B.

If you suspect problems because of shared cached devices, you can direct CPExpert to analyze the specific devices (using the SELECT option in DASGUIDE). CPExpert will then analyze only the devices selected. The analysis of shared DASD is accomplished you have specified **%LET SHARED = YES;** in USOURCE(DASGUIDE) to direct CPExpert to analyze shared DASD.

CPExpert performs the following processing if you have indicated that an analysis of potential conflicts between shared DASD should be performed:

- CPExpert determines whether the IOCP has indicated that the "worst" device selected for detailed analysis is attached to a control unit shared with another system[1]. If so, CPExpert performs an analysis of potential conflicts caused by shared DASD.

- CPExpert identifies other systems which reference the "worst" device. This identification is accomplished by analyzing the SMF Type 74 data in the performance data base relating to all other systems. The SMF Type 74 data contain the VOLSER for each device referenced. CPExpert simply selects SMF Type 74 information for the systems which reference the VOLSER of the "worst" device. This information is retained for more detailed analysis about potential conflicts.

  There is a potential problem with this method of identifying devices shared between systems. Multiple systems in the performance data base could use the same VOLSER to identify different devices. This could happen if the devices were not shared between systems.

  For example, suppose that CPExpert had identified PAGE01 as the "worst" device. Several system in the performance data base could reference VOLSER PAGE01, but the devices with VOLSER PAGE01 could be unique to each system. CPExpert would assume that all references by other systems to PAGE01 applied to the "worst" device being analyzed. The references could apply to a totally different device, and the other systems might not even share DASD with the system being analyzed.

  If this should be a problem (that is, if the DASD Component reports shared DASD conflicts with systems which do not share the device being analyzed), simply ignore the analysis produced by CPExpert[2].

[1]CPExpert tests whether the SHARED keyword of the CNTLUNIT macro in the IOCP was specified as either SHARED=Y or SHARED=YB.

[2]We do not feel that this problem will be common. It is described only to alert you to a potential incorrect analysis. If any user encounters this problem and it becomes annoying, code can be implemented to allow users to identify specific systems which share DASD with the system being analyzed. At present, this option seems to add unnecessary complexity to the user options.

- Once CPExpert has identified all systems which reference the "worst" device, CPExpert analyzes the DASD I/O characteristics of these systems with respect to the "worst" device. As described earlier, the analysis makes a basic assumption that the I/O activity from the different systems is random among the systems (that is, the I/O activity of System B is independent from the I/O activity of System A).

  If the I/O activity is random, CPExpert can conclude that delays on System A are caused by shared DASD conflicts only if (1) System B experienced the same delays with the "worst" device (in the case of PEND delays) or (2) System B generated a high I/O activity to the "worst" device.

  CPExpert can conclude that delays on System A are not caused by shared DASD conflicts if the data from System B does not exhibit these characteristics. The earlier discussion explains the rationale for these conclusions.

CPExpert will list statistics relating to potential conflicts, by system, by volume, and by RMF measurement interval. Additionally, CPExpert will identify the major cause of performance problems with the shared volume (i.e., seeking, PEND time, missed RPS delays, or missed cache hits).

Suggestion: You should use the information displayed by Rule DAS300 to assess the significance of the performance problems caused by shared DASD.

- You may conclude that the performance delays are acceptable, in which case you may take no action.

- You may conclude that the performance delays are not acceptable, in which case you may wish to take actions such as:

  - Schedule conflicting applications so that they do not access the volume at the same time.

  - Duplicate data or files and placing the duplicate information onto different volumes. This may be feasible for read-only data.

  - Move conflicting files to another shared volume.

Rule DAS310: Seeking probably was caused by shared DASD conflicts

Finding: CPExpert believes that the seeking performance problems were caused by shared DASD conflicts.

Impact: This finding is used to assess whether sharing DASD between systems or MVS images caused performance problems.

Logic flow: The following rules cause this rule to be invoked:

- DAS100: Volume with the worst overall performance
- DAS110: Seeking was the major cause of I/O response delay
- DAS300: Shared DASD caused performance problems

Discussion: If CPExpert determines that seeking was the major cause of I/O response delay and if the device is shared, CPExpert analyzes other systems in the performance data base which share the volume.

Seek delays can occur if the arm of the device has been moved by System B when System A attempts to access a cylinder. In this case, the DASD I/O operation from System A must move the arm to the desired cylinder, as a SEEK operation.

CPExpert computes the time required to perform SEEKs on System A. If the computed SEEK time is a major cause of performance problems, CPExpert analyzes the data from System B to determine whether System B generates a large number of I/O operations to the device.

If System B does **not** generate a relatively large number of I/O operations to the device, CPExpert concludes that there is **not** a conflict. If System B **does** generate a relatively large number of I/O operations to the device, CPExpert concludes that there **is** a conflict caused by sharing the device.

- There is little doubt about the validity of the first conclusion: if System B does not use the device, System B clearly cannot cause seek problems for System A.

- The second conclusion is based on an assumption: a large number of I/O operations from System B to the shared device will cause the seek problems for System A.

To be absolutely correct, CPExpert should process the configuration definitions for System B, process System B's channel and device information, and compute seek information for System B. CPExpert could then determine whether System B also experienced a high seek rate for the device. If both System A and System B experienced a high seek rate, CPExpert could be absolutely sure that there was a shared DASD conflict. This approach would unnecessarily use system resources and would be cumbersome to implement.

Consequently, CPExpert makes the assumption that I/O operations to the device are random between System A and System B. CPExpert thus can conclude that if System A experiences a high seek rate and System B significantly uses the device (exhibited by a high I/O rate), then System B must also experience a high seek rate. To assume otherwise would require that I/O from System B be coordinated with the I/O from System A, such that System B does not experience seeking similar to System A.

Since both System A experiences a high seek rate and System B experience a high I/O rate, CPExpert concludes that there is a conflict caused by the shared DASD.

Suggestion: You should use the information displayed by Rule DAS300 to assess the significance of the performance problems caused by shared DASD. Please refer to the suggestions associated with Rule DAS300 for alternative actions you may consider.

Rule DAS330: Large PEND time probably was caused by shared DASD conflicts

Finding: CPExpert believes that the large PEND time performance problems were caused by shared DASD conflicts.

Impact: This finding is used to assess whether sharing DASD between systems or MVS images caused performance problems.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS130: Major cause of I/O delay was PEND time
DAS132: Large PEND time may be caused by other system
DAS300: Shared DASD conflicts caused performance problems

Discussion: If CPExpert determines that large PEND time was the major cause of I/O response delay and if the device is shared, CPExpert analyzes other systems in the performance data base which share the volume.

The PEND time is a measure of the time an I/O operation waited (1) for a channel path to become available, (2) for the controller to become available, or (3) for the device because the device was busy to another system.

- If the I/O operation waited for a channel path, the I/O delay was caused by activity on the system issuing the I/O operation. Since channel paths are not shared between systems, the delay must occur at the controller or device level if the delay was caused by sharing DASD.

- If the I/O operation waited for the controller to become available, the I/O delay could be caused by activity on the system issuing the I/O operation or could be caused by another system which was using the controller.

- If the I/O operation waited for the device, the I/O delay was caused by another system which was using the device. The I/O Supervisor (IOS) on the system issuing the I/O operation could not know that the device was busy with another system (the device Unit Control Block would not reflect the "busy" status). Consequently, the IOS would believe the device was available and would issue the I/O operation. The I/O operation would wait while the device was busy to the other system. This time would be reflected as PEND time.

CPExpert analyzes the I/O activity from other systems sharing the device. The I/O activity measures used by CPExpert consists of the DISC time and CONN time from other systems. CPExpert substitutes the total RESERVE time for other systems if the total RESERVE time for other systems is greater than the I/O activity time.

CPExpert concludes the high PEND time is caused by other systems sharing the device if the device I/O activity (or total RESERVE time) is more than 25% of the PEND time experienced by the system being analyzed.

Suggestion: You should use the information displayed by Rule DAS300 to assess the significance of the performance problems caused by shared DASD. Please refer to the suggestions associated with Rule DAS300 for alternative actions you may consider.

Rule DAS360: Missed cache hits probably were caused by shared DASD conflicts

Finding: CPExpert believes that the missed cache hits performance problems were caused by shared DASD conflicts.

Impact: This finding is used to assess whether sharing DASD between systems or MVS images caused performance problems.

Logic flow: The following rules cause this rule to be invoked :

DAS100: Volume with the worst overall performance
DAS160: Missed cache read hits was major cause of I/O delay
DAS300: Shared DASD conflicts caused performance problems

Discussion: If CPExpert determines that cache missed read hits was the major cause of I/O response delay and if the device is shared, CPExpert analyzes other systems in the performance data base which share the volume.

Cache read hits can be caused by another system if the other system reads data from devices attached to the cached controller. If the data required by the other system is not in cache, the data must be brought into the cache to replace data already in the cache. The data being replaced might subsequently be required by the system being analyzed, and thus missed cache read hits would occur.

- If the data required by the other system is already in the cache, the data transfer will be reflected as CONN time by the other system.

- If the data must be brought into the cache, the mechanical time for the device will be reflected in DISC time. This is the time to position the arm (seek time) and to rotate the disk to the proper sector (latency). Rule DAS110 and Rule DAS310 would be produced if CPExpert estimated that seeking was a problem. Consequently, the residual time examined by Rule DAS360 is the latency delay.

In either of the above situations with cached shared devices, CPExpert analyzes the amount of I/O operations from System A and System B to the cached controller.

- If System B does **not** generate a relatively large number of I/O operations to the device, CPExpert concludes that there is **not** a conflict.

- If System B **does** generate a relatively large number of I/O operations to the device, CPExpert concludes that there **is** a conflict caused by sharing the device. The rationale for these conclusions is the same as was discussed in Rule DAS310 (seeking probably was caused by shared DASD conflicts).

The standard analysis performed by CPExpert may not detect a cache problem under two possible scenarios.

- If only one cached controller is attached to System A, CPExpert may not detect a problem with the device. This is because the logic employed by CPExpert selects devices with the most performance improvement within each **type of device** and then selects the **overall** "worst" devices for detailed analysis.

  CPExpert considers cached devices to be a unique type. If all devices on the cached controller received bad service caused by shared cached problems, CPExpert may not detect a performance problem with the cached devices. This is because all devices in the "device type" could have roughly equal poor service and no device would be **significantly** worse than the other devices in the device type. Consequently, CPExpert might not select any of the cached devices for detailed analysis

  If there are multiple cached controllers, there would be a larger number of "candidate" devices, and the standard analysis performed by CPExpert is more likely to identify any problem caused by shared cache controllers.

- The analysis performed by CPExpert may not detect a cache problem if the cache is being replaced with data from another volume. It is possible that System B could cause data from another volume to be loaded into the controller's cache. This volume could be a volume not be flagged as the "worst" performing volume when analyzing performance from the perspective of System A. From System A's perspective, accesses to the worst volume would simply not find required data in cache.

  Without analyzing the IOCP information for all systems, CPExpert cannot determine which volumes are attached to which controllers. Consequently, CPExpert cannot at present relate (1) poor performance for one volume on System A and (2) I/O operations to a different volume by System B.

If you suspect problems because of shared cached devices, you can direct CPExpert to analyze the specific devices (using the SELECT option in DASGUIDE). CPExpert will then analyze only the devices selected.

Suggestion: You should use the information displayed by Rule DAS300 to assess the significance of the performance problems caused by shared DASD. Please refer to the suggestions associated with Rule DAS300 for alternative actions you may consider.

Rule DAS380: Applications potentially causing shared DASD conflicts

Finding: CPExpert lists the applications which might cause shared DASD conflicts.

Impact: This finding is used to assess whether sharing DASD between systems or MVS images caused performance problems, and to identify specific applications which might cause conflicts.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS300: Shared DASD caused performance problems

This finding is produced only if the CPExpert modification to MXG or MICS is installed, so that SMF Type 30(DD) information is available.

Discussion: If CPExpert determines that there may be DASD conflicts caused by sharing DASD between systems of MVS images, and if the CPExpert modification to MXG or MICS is installed to collect application/device data, CPExpert performs further analysis.

CPExpert analyzes the SMF Type 74 information to acquire the device numbers associated with each VOLSER on every system. This processing is required because, while all systems would refer to a particular device by the VOLSER of the device, each system could reference the device using a different device number.

CPExpert creates the DASD30DD data set from the Type 30 records, using the modification to MXG or MICS. VOLSER information is not retained in the SMF Type 30 records; device numbers are retained in the SMF Type 30 records. Consequently, CPExpert must associate the unique device number used by each system, to the VOLSER of the device.

CPExpert then analyzes the DASD30DD data set created by the modification to MXG or MICS. CPExpert selects all job steps executing on **other** systems which referenced the device number of the device with the worst performance. For example, if CPExpert is analyzing System A, CPExpert would select all job steps executing on System B, System C, etc. if those job steps referenced the device with the worst performance.

CPExpert includes only references for time intervals when earlier analysis had concluded that shared DASD conflicts might cause a performance problem.

CPExpert produces a list of applications (listed by SMF Type 30 interval start and end times), ranked descendingly by I/O operations to the device with the worst performance.

Suggestion: Rule DAS380 is provided so you can identify the applications potentially causing shared DASD conflicts. You may wish to take the following actions:

- Reschedule the applications which cause shared DASD conflicts to a different processing time. The shared DASD conflicts may be eliminated or reduced if the applications can be rescheduled.

- Review the data sets on the volume to determine whether the data sets can be moved to other shared volumes (or to non-shared volumes). The shared DASD conflicts may be reduced or eliminated if the **data sets** are not required by systems sharing the volumes.

- Modify the applications causing shared DASD conflicts to reduce or eliminate their use of the data on the shared volume.

Finding: CPExpert lists the applications which might cause shared DASD conflicts during the RMF measurement intervals with the worst performance.

Impact: This finding is used to assess whether sharing DASD between systems or MVS images caused performance problems, and to identify specific applications which might cause the worst conflicts.

Logic flow: The following rules cause this rule to be invoked:

DAS100: Volume with the worst overall performance
DAS300: Shared DASD conflicts caused performance problems

This finding is produced only if the CPExpert modification to MXG or MICS is installed, so that SMF Type 30(DD) information is available.

Discussion: If CPExpert determines that there may be DASD conflicts caused by sharing DASD between systems of MVS images, and if the CPExpert modification to MXG or MICS is installed to collect application/device data, CPExpert performs further analysis. This analysis is described in Rule DAS380.

CPExpert further processes the information selected from the DASD30DD data set, to select job steps which referenced the device with the worst performance, **during the RMF measurement interval in which the device had the worst performance**. CPExpert includes only references for time intervals when earlier analysis had concluded that shared DASD conflicts might cause a performance problem.

CPExpert produces a list of applications (listed by SMF Type 30 interval start and end times), ranked descendingly by the access to the device with the worst performance. The SMF Type 30 interval data are prorated, if necessary, to reflect the approximate number of I/O operations during the worst RMF interval. Please refer to Section 5 (Chapter 2.1) for a discussion of prorating issues.

Suggestion: Rule DAS385 is provided so you can identify the applications potentially causing the worst shared DASD conflicts. You may wish to take the following actions:

- Reschedule the applications which cause shared DASD conflicts to a different processing time. The shared DASD conflicts may be eliminated or reduced if the applications can be rescheduled.

- Review the data sets on the volume to determine whether the data sets can be moved to other shared volumes (or to non-shared volumes). The shared DASD conflicts may be reduced or eliminated if the **data sets** are not required by systems sharing the volumes.

- Modify the applications causing shared DASD conflicts to reduce or eliminate their use of the data on the shared volume.

Rule DAS390: Performance problems apparently were not caused by shared DASD conflicts

Finding: CPExpert did not detect performance problems which were caused by shared DASD conflicts.

Impact: This finding is used to assess whether sharing DASD between systems or MVS images caused performance problems.

Logic flow: This is a basic finding. There are no predecessor rules.

Discussion: Parameters with the IOCP macro for the system indicated that the device with the worst performance was shared between systems or MVS images.

CPExpert analyzed data from all systems in the performance data base. The analysis did not indicate that the performance problems were caused by sharing the device.

This finding could result because the device was not configured on-line to another system, even though the IOCP indicated that the device was shared.

Alternatively, the other system might not have generated a large amount of I/O activity to the device.

In either case, CPExpert does not believe that sharing DASD caused conflicts.

Suggestion: There are no suggestions with the finding. Rule DAS390 is provided simply for information purposes.

Finding: CPExpert identifies the access characteristics of data sets managed by DFSMS (and reported in SMF Type 42 records), for the volumes that have the most potential for performance improvement.

Impact: This finding is used to assess the data sets that reside on the volumes with significant performance improvement potential.

Logic flow: This is a basic finding. There are no predecessor rules.

Discussion: If SMF Type 42 (Data Set Statistics) information is available[1] in a MXG performance data base, the DASD Component will process the MXG TYPE42DS file. The DASD Component select information describing data sets that were referenced during RMF measurement intervals in which the poorly performing device(s) exceeded the average performance. The result from this analysis is displayed in Rule DAS400.

There are several considerations with respect to the information produced by Rule DAS400:

- There can be many data sets referenced on the devices (hundreds or even thousands of data sets can be referenced). It is not helpful to have a large number of data sets listed. Consequently, CPExpert provides two guidance variables to control how many data sets are listed by Rule DAS400:

  - The **LIST42DS** variable can be used to limit the number of data sets listed in any RMF measurement interval, for a particular device. Only the number of data sets specified by the LIST42DS variable will be listed individually, and information about any remaining data sets will be summarized and listed on a single line.

    The default value for the LIST42DS variable is 10, indicating that 10 data sets will be listed for each RMF interval in which a poorly-performing device had a performance problem. You can alter this number of data sets listed by using the LIST42DS guidance variable.

  - Some data sets might have a low I/O activity and would not be interesting to analyze. CPExpert provides the **MIN42PCT** variable to limit the number of data sets listed in any RMF measurement interval,

[1] %LET TYPE42DS = Y; was specified in USOURCE(GENGUIDE).

for a particular device. Only data sets having a percent of activity for the total volume greater than the percent specified by the MIN42PCT variable will be listed individually, and any remaining data sets will be summarized and listed on a single line.

The default value for the MIN42PCT variable is 0.1, indicating that data sets will not be listed individually for any RMF interval unless the data set intensity of access (I/O rate * response time) was greater than 0.1% of the total volume intensity of access. You can alter this percent of data sets listed by using the MIN42PCT guidance variable.

For example, if you specified %LET MIN42PCT = 25, a maximum of 4 data sets would be listed in any RMF interval since no more than 4 data sets could have 25% or higher access intensity.

Please note that, regardless of the data set access intensity of any particular data set, only the number of data sets specified by the **LIST42DS** guidance variable (described earlier) will be listed. This means that there are two ways to limit the number of data sets listed: (1) the LIST42DS which limits the number of data sets listed in any RMF interval, and (2) the **MIN42PCT** guidance variable which limits the data sets listed to those that exceed the specified percent.

- Some devices might have data sets that are not managed by DFSMS. In this case, information would not be available in SMF Type 42 records and the data sets could not be reported by CPExpert.

- Depending on the timing in which SMF writes Type 42 records, some information might not be represented in the RMF Type 74 intervals by which the DASD Component analyzes DASD performance problems. Consequently, the information presented in Rule DAS400 might not correspond to the information presented in other rules (for example, might not correspond to information presented in Rule DAS100).

- SMF writes Type 42 (Data Set Statistics) records when (1) a DASD data set is closed, or (2) immediately after the recording of the SMF Type 30 interval record. There are two implications of this SMF write characteristic:

  - A data set likely would be closed at some time between SMF writing Type 30 interval records. Consequently, some Type 42 (Data Set Statistics) records would not correspond to the SMF Type 30 interval, and the data set information reported by the DASD Component might not correspond to the information reported for the device by other rules.

  - The SMF Type 30 recording interval controls when Type 42 (Data Set Statistics) records are written (for those data sets that are OPEN when the Type 30 record is written). However, the SMF Type 30 recording interval might not correspond with the RMF recording interval. If the SMF and RMF writing is synchronized, the SMF Type 30 recording interval would be synchronized with the RMF recording interval. However, the synchronization option might not have been selected.

Even if the synchronization option had been selected, a different recording interval frequency could have been selected for Type 30 and RMF recording. For example, RMF recording could occur at 15 minute intervals, but SMF Type 30 recording could occur at 30 minute intervals. While the intervals could be synchronized, there would be twice as many RMF recording intervals as there were SMF Type 30 (and Type 42) recording intervals.

The SMF and RMF recording considerations might cause consternation if a user were not aware of the potential problems. If the SMF and RMF recording intervals are synchronized and are for the same duration, and if data sets are OPEN for the duration of the recording interval, information presented in Rule DAS400 will correspond well with device-related information (and application-related information) presented by other rules. If these conditions are not present, the information presented by Rule DAS400 might not correspond well with information presented elsewhere by the DASD Component.

Please send a note to Don_Deese@cpexpert.com if these problems become troublesome. Perhaps the logic can be improved (by prorating the DASD42DS information, for example).

Suggestion: You should use the information displayed by Rule DAS400 to assess the data access characteristics for DFSMS-managed data sets residing on the device(s) with the most potential for performance improvement.

Finding: CPExpert determined that an excessive number of Control Area (CA) splits occurred for the VSAM data sets listed.

Impact: This finding is used to assess problems or potential problems with VSAM Control Area splits for the data sets listed.

Discussion: A VSAM file structure consists of one or more *Control Intervals (CIs)* and one or more *Control Areas (CAs)*.

- A **Control Interval** is a continuous area of direct access storage that VSAM uses to store logical records. When a logical record is read from direct access storage, the entire Control Interval containing the record is read into a VSAM buffer in virtual storage. The desired logical record is then transferred from the VSAM buffer to a user-defined buffer or work area. While logical records within a Control Interval may vary in length, all Control Intervals in a specific VSAM data set are of the same length.

  In addition to the logical records, a Control Interval consists of free space, and control information. The free space initially is unused, and is used to accommodate inserted logical records or for changes in the length of logical records. The control information describes the amount and location of free space, and describes the length of records and how many adjacent records are of the same length.

- A **Control Area** contains one or more Control Intervals. The Control Intervals are grouped together into fixed-length contiguous areas of direct access storage. A VSAM data set consists of one or more Control Areas.

  The minimum size of a Control Area is one track of DASD storage, and the maximum size of a Control Area is one cylinder. The size of a Control Area is not specified by a user; the size of the CA is *calculated* by VSAM based on the amount of space allocated to a VSAM data set.

Logical records can be inserted into a VSAM keyed sequenced data set (KSDS) or a variable relative record data set (VRRDS), or a record length can be increased. When either of the actions occur, space must be available in a Control Interval to accommodate the new record (or accommodate the increased length). The space for the new record (or increased length) normally is obtained from the *free space* that was allocated to the Control Interval[1] when the data set was loaded.

If enough free space is not available in the Control Interval, a *Control Interval Split* occurs. When a Control Interval is split, approximately one-half of the logical records in the original Control Interval will be transferred to a new Control Interval, and the records are then deleted in the old Control Interval. Space for the new Control Interval is obtained from the free space that is associated with the Control Area. Thus, inserted records or longer records are inserted into Control Intervals using the free space associated with the Control Interval. If the Control Interval does not have sufficient space, a new Control Interval will be created, using free space in the Control Area[2] to which the original Control Interval belongs.

If enough free space is not available in the Control Area, a *Control Area Split* occurs. When a Control Area is split, approximately one-half of the Control Intervals from the Control Area are moved to the end of the data set. The records are then deleted from the old Control Area (requiring additional I/O operations). For example, if the CA is one cylinder, then approximately one-half of the records in the cylinder is moved to the end of the data set, and the records that were moved re then deleted from their location in the original cylinder. Depending on the size of the Control Area, considerable overhead[3] can result from Control Area splits.

Additionally, locking may be involved during Control Area split processing, depending on the SHARE options associated with the VSAM data set. The below table shows the level at which locking occurs, and the condition that causes the lock level.

LOCK LEVEL	**CONDITION**
Control Interval	Adding a record or updating a record in place, without causing a Control Interval split
Control Area	Adding a record or updating a record in place, causing a Control Interval split, but not causing a Control Area split.
Data Set	Adding a record or updating a record in place, causing a Control Interval split, and causing a Control Area split.

[1]The amount of free space reserved for each Control Interval is controlled by the FREESPACE parameter. The first value of the FREESPACE keyword specifies a percent of each Control Interval that is to be reserved for free space.

[2]The amount of free space reserved for each Control Area is controlled by the FREESPACE parameter. The second value of the FREESPACE parameter specifies a percent of each Control Area that is to be reserved for free space.

[3]Note that the overhead occurs only during the CA split. There is essentially no additional overhead after the CA split.

Considering the overhead involved and the locking that can occur, Control Area splits should be avoided if possible.

CPExpert examines the SMF Type 64 information contained in MXG TYPE64 data set to identify VSAM data sets that have excessive Control Area splits.

CPExpert sums the ACCASPLT variable (the number of CA splits since the data set was created) and the CASPLITS variable (the number of CA splits with the current OPEN of the data set). CPExpert compares this sum with the **CASPLITS** guidance variable in USOURCE(DASGUIDE). CPExpert produces Rule DAS600 when the total number of CA splits exceeds the value specified by the **CASPLITS** guidance variable.

The default value for the **CASPLITS** guidance variable is 10, indicating that CPExpert should produce Rule DAS600 when a VSAM data set experienced more than 10 CA splits.

The following example illustrates the output from Rule DAS600:

```
RULE DAS600:  EXCESSIVE CONTROL AREA SPLITS OCCURRED

   VOLSER: RLS014.  More than 10 Control Area (CA) splits occurred for the
   VSAM data sets listed below.  CA splits cause considerable overhead
   during the split, and CA splits should be avoided.
                                                             ..       TOTAL CA  CA SPLITS
   SMF TIME STAMP    JOB NAME  VSAM DATA SET                 ..        SPLITS   THIS OPEN
   10:30,29AUG2001   CICS2AGC  RLSADSW.VF03D.DATAENDB.DATA.................   115         9
   10:30,29AUG2001   CICS2AGC  RLSADSW.VF04D.DATAENDB.DATA.................   117        11
   10:30,29AUG2001   CICS2AGA  RLSADSW.VF07D.ITEMACT.DATA..................    63        49
```

Although not shown in this example, CPExpert also reports the total number of inserts to the VSAM data set and inserts for the current OPEN of the VSAM data set.

Suggestion: The action taken to reduce CA splits requires a reorganization of the data set, often with different values for the FREESPACE parameter. If CA splits occur frequently for a VSAM data set, you should consider the following alternatives:

- **Reduce CI free space and increase CA free space**. Unless inserts (or increases in record length) are evenly distributed across Control Intervals, you should reduce (or eliminate) the Control Interval FREESPACE amount and increase the Control Area FREESPACE amount. Control Interval splits do not cause much overhead (little data is moved and the movement remains within the Control Area). Reducing or eliminating the Control Interval FREESPACE amount eliminates wasted DASD space, unless inserts take place for most Control Intervals. Adding free space to the Control Area will often significantly reduce the number of Control Area splits and significantly reduce the overhead associated with the Control Area splits.

If new records will be evenly distributed throughout the data set, control area free space should equal the percentage of records to be added to the data set after the data set is loaded (specify FREESPACE (0 nn), where nn equals the percentage of records to be added.)

Unfortunately, simply increasing the amount of Control Area free space might not be a good solution. Specifying too much free space can result in more direct access storage required to contain the data set and much of this space might be wasted if inserts are clustered.

- **Allow the insert pattern to control where splits occur**. If insertions will be unevenly distributed throughout the data set and you cannot tell where the insertions are likely to occur, you can specify a small amount of free space. In this case, splits will occur *in those areas where insertions occur*. This option reduces the amount of wasted DASD space caused by a large amount of unused free space.

- **Allocate CA free space where clustering occurs**. If you can determine the clustering nature of the record insertions and the data set is large, you might consider loading the data set in stages[4], with free space specified in areas where clustering is expected.

- **Spread the overhead**. If the VSAM data set is used **primarily by on-line applications**, you might specify a relatively small primary allocation, so that the impact of Control Area splits would be spread across the transactions encountering the Control Area splits. If a Control Area were one cylinder, the considerable overhead of splitting the Control Area would be incurred by the transaction that caused a Control Area split to occur. This might be a relatively random event (depending on the characteristics of other transactions and which user submitted the transactions), and would result in unpredictably poor response time. One way to reduce the impact on any particular on-line user would be to make the Control Area relatively small (for example, use a *primary allocation unit* of only a track, which would cause the Control Area to be only a track) so that when a CA split did occur, the immediate impact would not be as significant.

[4] IBM shows an example of this approach in *DFSMS: Using Data Sets* (Section 2.5.3.2: Altering the Free Space Specification When Loading a Data Set)

- **You can not determine the clustering nature of the record insertions and the data set is relatively small**. If this is the situation, you probably should ignore the CA splits reported by DAS600, since the CA splits should not continue to occur. If the CA splits **should** continue to occur, you probably will see the data set allocation begin to acquire additional extents. This is because the size of the CA is determined by VSAM based on the primary and secondary allocation. In this situation, CPExpert will produce Rule DAS604 (Excessive secondary extents were allocated).

- If none of the above actions are appropriate, you can change the **CASPLITS** guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the CASPLITS guidance variable if you feel that Rule DAS600 is produced too often, or if you do not wish to take action when only 10 CA splits occur.

- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.11.4.2.3: Control Interval Splits
Section 2.5.2: Optimizing Control Area Size
Section 2.5.3.1: Selecting the Optimal Percentage of Free Space
Section 2.5.3.2: Altering the Free Space Specification When Loading a Data Set

Rule DAS604: Excessive secondary extents were allocated

Finding: CPExpert determined that an excessive number of secondary extents were allocated for the keyed sequenced data set (KSDS) or variable relative record data set (VRRDS) VSAM data sets listed.

Impact: This finding is used to assess problems or potential problems with VSAM primary and secondary allocation values for the data sets listed. The impact can be significant, particularly with on-line applications.

Discussion: When a VSAM data set is allocated, space allocation amounts normally are specified for both a primary allocation and a secondary allocation. When the primary amount on the first volume is used up, a secondary amount is allocated on that volume by the end-of-volume (EOV) routine, using the amount specified for the secondary allocation.

This space allocation process can be repeated until the volume is out of space or until the extent limit is reached. Depending on the type of data set allocation request, a new volume may be used if the current volume is out of space.

A large number of I/O operations is involved when the secondary allocation takes place. Consequently, a small amount of space should not be specified for the primary or secondary allocation value, especially for a KSDS data set or for a VRRDS data set[1].

CPExpert examines the SMF Type 64 information contained in MXG TYPE64 data set to identify VSAM KSDS or VRRDS data sets that have excessive secondary allocations.

CPExpert compares NREXTNTS variable (the number of secondary extents in the VSAM data set this OPEN) with the **EXTENTS** guidance variable in USOURCE(DASGUIDE). CPExpert produces Rule DAS604 when the NREXTENT (the total number of extents) is greater than one, and the number of secondary extents allocated for this OPEN exceeds the value specified by the EXTENTS guidance variable.

The default value for the EXTENTS guidance variable is 0, indicating that CPExpert should produce Rule DAS604 when any secondary extent was allocated for the VSAM data sets listed.

---

[1] If the VSAM data set is used **primarily by on-line applications**, you might establish a relatively small primary allocation (which would result in a small Control Area), so the impact of Control Area splits would be spread across the transactions encountering the Control Area splits. See Rule DAS600 for further discussion about this issue.

The following example illustrates the output from Rule DAS604:

```
RULE DAS604:  EXCESSIVE SECONDARY EXTENTS WERE ALLOCATED

   VOLSER: RLS014.  More than 0 secondary extents were allocated for the VSAM
   data sets listed below.  There are a large number of I/O operations
   involved when the secondary allocation takes place.  Depending on
   your I/O configuration and constraints, a large number of secondary
   allocations could cause significant performance degradation of
   applications.  A large number of secondary extents normally means
   that the secondary allocation value is too small.  This problem can
   be particularly serious for KSDS or VRRDS data sets.
                                                                     ..        TOTAL    EXTENTS    FILE
   SMF TIME STAMP    JOB NAME  VSAM DATA SET                         ..       EXTENTS  THIS OPEN   TYPE
   10:00,29AUG2000   CICS2AGC  RLSADSW.VF03D.DATAENDB.DATA.................      10        6       KSDS DATA
   10:30,29AUG2000   CICS2AGC  RLSADSW.VF03D.DATAENDB.DATA.................      10        6       KSDS DATA
```

Suggestion: If CPExpert produces Rule DAS604, you should consider the following alternatives:

- If a relatively large number of secondary allocations is made, you should consider increasing the primary allocation amount for the data set.
- Additionally, you should consider increasing the secondary allocation amount for the data set.
- If the above actions are not appropriate, you can change the EXTENTS guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the EXTENTS guidance variable if you feel that Rule DAS604 is produced too often, or if you do not wish to take action when secondary extents are allocated.
- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.2.2.4: Allocating Space for VSAM Data Sets

VSAM Demystified Redbook (SG24-6105)
Section 2.6: Parameters affecting performance

Finding: CPExpert determined that an excessive number of secondary extents were allocated for the keyed sequenced data set (KSDS) or a variable relative record data set (VRRDS) VSAM data sets listed.

Impact: This finding is used to assess problems or potential problems with VSAM primary and secondary allocation values for the data sets listed. The impact can be significant, particularly with on-line applications.

Discussion: When a VSAM data set is allocated, space allocation amounts normally are specified for both a primary allocation and a secondary allocation. When the primary amount on the first volume is used up, a secondary amount is allocated on that volume by the end-of-volume (EOV) routine, using the amount specified for the secondary allocation.

This space allocation process can be repeated until the volume is out of space or until the extent limit is reached. Depending on the type of data set allocation request, a new volume may be used if the current volume is out of space.

A large number of I/O operations is involved when the secondary allocation takes place. Consequently, a small amount of space should not be specified for the primary or secondary allocation value, especially for a KSDS data set or for a VRRDS data set[1].

CPExpert examines the SMF Type 64 information contained in MXG TYPE64 data set to identify VSAM KSDS or VRRDS data sets that have excessive secondary allocations.

CPExpert compares NREXTNTS variable (the number of secondary extents in the VSAM data set this OPEN) with the **EXTENTS** guidance variable in USOURCE(DASGUIDE). CPExpert produces Rule DAS605 when the NREXTENT (the total number of extents) is greater than one, and the number of secondary extents allocated for the current OPEN exceeds the value specified by the EXTENTS guidance variable.

When Rule DAS604 is produced, CPExpert analyzes the primary and secondary allocation units. If the primary or secondary allocation unit is in tracks, CPExpert produces Rule DAS605 to reflect the allocation values.

---

[1] If the VSAM data set is used **primarily by on-line applications**, you might establish a relatively small primary allocation (which would result in a small Control Area), so the impact of Control Area splits would be spread across the transactions encountering the Control Area splits. See Rule DAS600 for further discussion about this issue.

The following example illustrates the output from Rule DAS605:

```
RULE DAS605:  PRIMARY OR SECONDARY ALLOCATION UNIT WAS SMALL

   VOLSER: RLS01C.  More than 0 secondary extents were allocated for the
   VSAM data sets listed below.  Either the primary allocation or
   the secondary allocation (or both) was smaller than one cylinder.
   In addition to causing multiple extents, this allocation size also
   means that the Control Area (CA) size is less than one cylinder.
   You should consider increasing the allocation units to use cylinders
   rather than tracks. The below shows the number of tracks used for the
   space allocation, and the number of extents:
                                                                   ..         PRIMARY   -----SECONDARY-----
   SMF TIME STAMP    JOB NAME  VSAM DATA SET                       ..        ALLOCATION ALLOCATION  EXTENTS
   10:30,29AUG2000   CICS2ACA  RLSADSW.VF04D.DATAENDB.INDEX................    2 TRKS      1 TRK        3
   11:00,29AUG2000   CICS2ACA  RLSADSW.VF04D.DATAENDB.INDEX................    2 TRKS      1 TRK        3
```

Suggestion: If CPExpert produces Rule DAS605, you should consider the following alternatives:

- If a relatively large number of secondary allocations is made, you should consider increasing the primary allocation amount for the data set.
- Alternatively, you should consider increasing the secondary allocation amount for the data set.
- If the above actions are not appropriate, you can change the EXTENTS guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the EXTENTS guidance variable if you feel that Rule DAS600 is produced too often, or if you do not wish to take action when secondary extents are allocated.
- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.2.2.4: Allocating Space for VSAM Data Sets

VSAM Demystified Redbook (SG24-6105)
Section 2.6: Parameters affecting performance

Finding: CPExpert determined that an excessive number of secondary extents were allocated for the keyed sequenced data set (KSDS) or variable relative record data set (VRRDS) VSAM data sets listed, and either the primary or secondary allocation unit was small.

Impact: This finding is used to assess problems or potential problems with VSAM primary and secondary allocation values for the data sets listed. The impact can be significant, particularly with on-line applications.

Discussion: When a VSAM data set is allocated, space allocation amounts normally are specified for both a primary allocation and a secondary allocation. When the primary amount on the first volume is used up, a secondary amount is allocated on that volume by the end-of-volume (EOV) routine, using the amount specified for the secondary allocation.

This space allocation process can be repeated until the volume is out of space or until the extent limit is reached. Depending on the type of data set allocation request, a new volume may be used if the current volume is out of space.

A large number of I/O operations is involved when the secondary allocation takes place. Consequently, a small amount of space should not be specified for the primary or secondary allocation value, especially for a KSDS data set or for a VRRDS data set[1].

CPExpert examines the SMF Type 64 information contained in MXG TYPE64 data set to identify VSAM KSDS or VRRDS data sets that have excessive secondary allocations.

CPExpert compares NREXTNTS variable (the number of secondary extents in the VSAM data set this OPEN) with the **EXTENTS** guidance variable in USOURCE(DASGUIDE). CPExpert produces Rule DAS604 when the NREXTENT (the total number of extents) is greater than one, and the number of secondary extents allocated for this OPEN exceeds the value specified by the EXTENTS guidance variable.

---

[1]If the VSAM data set is used **primarily by on-line applications**, you might establish a relatively small primary allocation (which would result in a small Control Area), so the impact of Control Area splits would be spread across the transactions encountering the Control Area splits. See Rule DAS600 for further discussion about this issue.

When Rule DAS604 is produced, CPExpert analyzes the primary and secondary allocation units. If the primary or secondary allocation unit is in cylinders, CPExpert produces Rule DAS606 to reflect the allocation values.

The following example illustrates the output from Rule DAS606:

```
RULE DAS606:  PRIMARY OR SECONDARY ALLOCATION UNIT WAS SMALL

   VOLSER: RLS014.  More than 0 secondary extents were allocated for the
   VSAM data sets listed below, even though cylinders were used as the
   allocation unit.  The allocation specified for either the primary
   or secondary allocation should be increased.  The below shows the
   primary and secondary space allocation and number of extents:
                                                            ..         PRIMARY   -----SECONDARY-----
   SMF TIME STAMP    JOB NAME  VSAM DATA SET                ..        ALLOCATION ALLOCATION  EXTENTS
   10:00,29AUG2000   CICS2AGC  RLSADSW.VF03D.DATAENDB.DATA.................    30 CYLS    10 CYLS        9
   10:30,29AUG2000   CICS2AGC  RLSADSW.VF03D.DATAENDB.DATA.................    30 CYLS    10 CYLS        9
```

Suggestion: If CPExpert produces Rule DAS606, you should consider the following alternatives:

- If a relatively large number of secondary allocations were made, you should consider increasing the primary allocation amount for the data set.
- Alternatively, you should consider increasing the secondary allocation amount for the data set.
- If the above actions are not appropriate, you can change the EXTENTS guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the EXTENTS guidance variable if you feel that Rule DAS600 is produced too often, or if you do not wish to take action when excessive secondary extents are allocated.
- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.2.2.4: Allocating Space for VSAM Data Sets

VSAM Demystified Redbook (SG24-6105)
Section 2.6: Parameters affecting performance

Rule DAS607: VSAM data set is close to maximum number of extents

Finding: CPExpert determined that a VSAM data set has had a significant number of secondary allocations, and is close to reaching the maximum number of extents.

Impact: This finding is used to assess potential problems with VSAM data sets reaching the maximum number of extents. The finding can have a HIGH IMPACT if the VSAM data set reaches the maximum number of extents.

Discussion: When a VSAM data set is allocated, space allocation amounts normally are specified for both a primary allocation and a secondary allocation. When the primary amount on the first volume is used up, a secondary amount is allocated on that volume by the end-of-volume (EOV) routine, using the amount specified for the secondary allocation.

This space allocation process can be repeated until the volume is out of space or until the extent limit is reached. Depending on the type of data set allocation request, a new volume may be used if the current volume is out of space.

A VSAM data set can have up to 255 extents per component, and a striped VSAM data set can have up to 255 extents per stripe. The last four extents are reserved for extending a data set when the last extent cannot be allocated in one piece. (VSAM attempts to extend a data set when the total number of extents is less than 250.

When a multivolume VSAM data set extends to the next volume, the data class specifies if the initial space allocated on that volume is the primary or secondary amount. The default is the primary amount. After the primary amount of space is used up, space is allocated in secondary amounts. By using a data class, it is possible to indicate whether to take a primary or secondary amount when VSAM extends to a new volume.

An ABEND results when an extend as attempted, but the maximum number of extents was reached. CPExpert attempts to provide an “early warning” of this potential situation.

CPExpert examines the SMF Type 64 information contained in MXG TYPE64 data set to identify VSAM data sets that have used a significant number of extents, such that there is danger of reaching the maximum extents. CPExpert compares NREXTENT variable (the total number of extents in the VSAM data set) with the **MXEXTENT** guidance variable in USOURCE(DASGUIDE). CPExpert produces Rule DAS607 when the

NREXTENT is greater than the value specified by the MXEXTENT guidance variable **and** at least one extent was allocated during the current OPEN of the data set (CPExpert uses the NREXTNTS variable in TYPE64 for this decision).

The default value of the MXEXTENT guidance variable is 225, indicating that CPExpert should produce Rule DAS607 when at least 225 extents have been allocated for a VSAM data set. Since the maximum allowable is 255, the default value provides a threshold at which CPExpert provides notification that there is a potential problem.

The following example illustrates the output from Rule DAS607:

```
RULE DAS607: VSAM DATA SET IS CLOSE TO MAXIMUM NUMBER OF EXTENTS

   VOLSER: RLS003.  More than 225 extents were allocated for the VSAM data
   sets listed below.  The VSAM data sets are approaching the maximum
   number of extents allowed. The below shows the number of extents
   and the primary and secondary space allocation:
                                                             ..         TOTAL    EXTENTS    ---ALLOCATIONS---
   SMF TIME STAMP    JOB NAME  VSAM DATA SET                 ..        EXTENTS  THIS OPEN   PRIMARY SECONDARY
   10:30,11MAR2002   CICS2ABA  RLSADSW.VF01D.DATAENDB.DATA.................     229        4       30 CYL    100 CYL
```

Suggestion: If CPExpert produces Rule DAS607, you should consider the following alternatives:

- Determine (potentially by consulting with applications personnel) whether the VSAM data set is expected to continue increasing. If the VSAM data set is expected to continue increasing, you should take action. IBM suggests that you (1) use the access method services REPRO command to make a backup copy of the cluster that contains the data set, (2) delete the cluster from the catalog with the DELETE command, (3) use the DEFINE command to redefine the cluster in the catalog with increased space allocation, and (4) reload the backup of the cluster with the REPRO command. If this alternative is selected, you should consider **significantly** increasing the primary allocation value and consider increasing the secondary allocation value.

- Alternatively, you should examine the amount of space actually used in the VSAM data set by logical records, compared with free space or space used by deleted logical records (depending on the type of VSAM data set). It is possible that inappropriate values have been specified for the FREESPACE parameter.

  The percentages of free space should yield full records and full control intervals with a minimum amount of unusable space. If too much free space was specified for the Control Intervals or Control Areas, there would be more direct access storage required to contain the data set. This unnecessary free space could cause the data set to be unnecessarily large, and cause extents to be acquired more often.

You can use LISTCAT to examine statistics about the space used and the free space allocated and used.

- If the above actions are not appropriate, you can change the MXEXTENT guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the MXEXTENT guidance variable if you feel that Rule DAS607 is produced prematurely.

- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.2.2.4: Allocating Space for VSAM Data Sets
Section 2.5.3: Optimizing Free Space Distribution

VSAM Demystified Redbook (SG24-6105)
Section 2.6: Parameters affecting performance

Rule DAS610: Relatively small CI size was used for sequential processing

Finding: CPExpert noticed that a relatively small CI size was used for sequential processing. This finding applies only if SMF Type 42 (Data Set Statistics)[1] and SMF Type 64 (VSAM Statistics) records are available in a MXG performance data base.

Impact: This can have a LOW IMPACT, MEDIUM IMPACT, or HIGH IMPACT on the performance of applications referencing the VSAM data. The level of impact depends on the number of sequential I/O operations.

Discussion: A VSAM file structure consists of one or more *Control Intervals (CIs)* and one or more *Control Areas (CAs)*.

- A **Control Interval** is a continuous area of direct access storage that VSAM uses to store logical records. When a logical record is read from direct access storage, the entire Control Interval containing the record is read into a VSAM buffer in virtual storage. The desired logical record is then transferred from the VSAM buffer to a user-defined buffer or work area. While logical records within a Control Interval may vary in length, all Control Intervals in a specific VSAM data set are of the same length.

- A **Control Area** contains one or more Control Intervals. The Control Intervals are grouped together into fixed-length contiguous areas of direct access storage. A VSAM data set is composed of one or more Control Areas.

Each VSAM data set is defined as a cluster of one or more components.

- The *data component* is the part of a VSAM data set, alternate index, or catalog that contains the data records. The minimum size of a Control Area for a data component is one track, and the maximum size is one cylinder of DASD storage.

- The *index component* is a collection of logically sequenced keys. A key is a value taken from a fixed defined field in each logical record in the VSAM data set. The key identifies the record's position in the data set. Using the index, VSAM is able to randomly retrieve a record from the data component when a request is made for a record with a certain key. The size of the Control Area for an index component is one track of DASD storage.

[1] **%LET TYPE42DS = Y;** must be specified in USOURCE(GENGUIDE) must be specified in USOURCE(GENGUIDE) or USOURCE(DASGUIDE) to advise CPExpert that TYPE42DS is available.

The size of Control Intervals can vary from one VSAM data set to another, but all the Control Intervals of a particular data set component must be the same length. Control interval size affects record processing speed and storage requirements:

- Data sets with large control interval sizes require more buffer space in virtual storage.
- Data sets with large control interval sizes require fewer I/O operations to bring a given number of records into virtual storage, because fewer index records must be read.
- Free space is used more efficiently as control interval size increases relative to data record size. This is because there are fewer Control Interval splits and less wasted space.

The type of processing that is used should guide the choice of control interval size, particularly for the data component:

- **Sequential processing**. When sequential processing accounts for most of the accesses, a large Control Interval for the data component would normally be a good choice. This is because multiple records can be read into buffers and processed sequentially. For example, given a 16KB data buffer space, it is better to read two 8 KB Control Intervals with one I/O operation, than to read four 4 KB Control Intervals with two I/O operations.
- **Direct processing**. When direct processing accounts for most of the accesses, a small Control Interval for the data component is preferable. This is because only one logical record is retrieved at a time with direct access. Since a Control Interval normally contains several logical records, a large Control Interval would create unnecessary I/O overhead reading the logical records that would not be accessed. IBM suggests that a 4096 byte Control Interval normally would be appropriate for direct access.
- **Mixed processing**. If the processing is a mixture of sequential and direct, a small Control Interval for the data component with multiple buffers for sequential processing can be a good choice. The small Control Interval would reduce the I/O processing when using direct access to reference a logical record. The multiple buffers would allow read-ahead (and I/O efficiency) when using sequential processing.

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information to compute the percent of sequential accesses to the data component of the VSAM data set, using the following algorithm:

$$\textit{Percent sequential accesses} = \frac{S42AMSRB}{S42AMSRB + S42AMDRB}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces Rule DAS610 when the percent of sequential accesses for the data component was greater than the **PCTSEQ** guidance variable in USOURCE(DASGUIDE), and the Control Interval size was less than 8192 bytes.

The default value for the PCTSEQ guidance variable is 80%, so CPExpert will produce Rule DAS610 when more than 80% of the accesses were sequential for the data component, and the Control Interval size was less than 8192 bytes.

The following example illustrates the output from Rule DAS610:

```
RULE DAS610:  RELATIVELY SMALL CI SIZE USED FOR SEQUENTIAL PROCESSING

   VOLSER: D83NE2.  A relatively small data Control Interval (CI) size
   was used for the VSAM data sets listed below, yet more than 80% of
   the accesses were for sequential processing.  You normally should define
   a CI size of 8KB or larger for VSAM data sets used mostly for sequential
   processing.  The I/O RATE is for the time the data set was open.
                                                          ..       I/O    -ACCESS TYPE (PCT)-    CI
   SMF TIME STAMP    JOB NAME  VSAM DATA SET              ..       RATE   SEQUENTIAL   DIRECT   SIZE
   17:00,14OCT1997   NETAK31.  OPSYS.NV31.J90.AOC.DSILOGP.DATA............  19.2     100.0        0.0   4096
```

While not shown in the above example, CPExpert also shows the OPEN time for the VSAM data set. This normally is the duration of the current OPEN. If %LET VSAMSMRY=Y; was specified in USOURCE(DASGUIDE), the OPEN time represents the sum of the times the VSAM data was OPEN for all TYPE64 records in the performance data base.

Suggestion: If CPExpert produces Rule DAS610, you should consider the following alternatives:

- For sequential processing of logical records, system performance normally will be better if multiple logical records are read with a single I/O operation. Consequently, a larger Control Interval would normally improve performance since read-ahead I/O processing could be performed, and you should consider increasing the Control Interval.

- If the above action is not appropriate, you can change the **PCTSEQ** guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the PCTSEQ guidance variable if you feel that Rule DAS610 is produced prematurely.

- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.5.1.2: Data Control Interval Size

Rule DAS611: Relatively large CI size was used for direct processing

Finding: CPExpert noticed that a relatively large CI size was used for direct processing. This finding applies only if SMF Type 42 (Data Set Statistics)[1] and SMF Type 64 (VSAM Statistics) records are available in a MXG performance data base.

Impact: This can have a LOW IMPACT, MEDIUM IMPACT, or HIGH IMPACT on the performance of applications referencing the VSAM data. The level of impact depends on the number of direct I/O operations.

Discussion: A VSAM file structure consists of one or more *Control Intervals (CIs)* and one or more *Control Areas (CAs)*.

- A **Control Interval** is a continuous area of direct access storage that VSAM uses to store logical records. When a logical record is read from direct access storage, the entire Control Interval containing the record is read into a VSAM buffer in virtual storage. The desired logical record is then transferred from the VSAM buffer to a user-defined buffer or work area. While logical records within a Control Interval may vary in length, all Control Intervals in a specific VSAM data set are of the same length.

- A **Control Area** contains one or more Control Intervals. The Control Intervals are grouped together into fixed-length contiguous areas of direct access storage. A VSAM data set is composed of one or more control areas.

Each VSAM data set is defined as a cluster of one or more components.

- The *data component* is the part of a VSAM data set, alternate index, or catalog that contains the data records. The minimum size of a Control Area for a data component is one track, and the maximum size is one cylinder of DASD storage.

- The *index component* is a collection of logically sequenced keys. A key is a value taken from a fixed defined field in each logical record in the VSAM data set. The key identifies the record's position in the data set. Using the index, VSAM is able to randomly retrieve a record from the data component when a request is made for a record with a certain key. The size of the Control Area for an index component is one track of DASD storage.

---

[1] **%LET TYPE42DS = Y;** must be specified in USOURCE(GENGUIDE) must be specified in USOURCE(GENGUIDE) or USOURCE(DASGUIDE) to advise CPExpert that TYPE42DS is available.

The size of Control Intervals can vary from one VSAM data set to another, but all the Control Intervals of a particular data set component must be the same length. Control interval size affects record processing speed and storage requirements:

- Data sets with large control interval sizes require more buffer space in virtual storage.
- Data sets with large control interval sizes require fewer I/O operations to bring a given number of records into virtual storage, because fewer index records must be read.
- Free space is used more efficiently as control interval size increases relative to data record size. This is because there are fewer Control Interval splits and less wasted space.

The type of processing that is used should guide the choice of control interval size, particularly for the data component:

- **Sequential processing**. When sequential processing accounts for most of the accesses, a large Control Interval for the data component would normally be a good choice. This is because multiple records can be read into buffers and processed sequentially. For example, given a 16KB data buffer space, it is better to read two 8 KB Control Intervals with one I/O operation, than to read four 4 KB Control Intervals with two I/O operations.
- **Direct processing**. When direct processing accounts for most of the accesses, a small Control Interval for the data component is preferable. This is because only one logical record is retrieved at a time with direct access. Since a Control Interval normally contains several logical records, a large Control Interval would create unnecessary I/O overhead reading the logical records that would not be accessed. IBM suggests that a 4096 byte Control Interval normally would be appropriate for direct access.
- **Mixed processing**. If the processing is a mixture of sequential and direct, a small Control Interval for the data component with multiple buffers for sequential processing can be a good choice. The small Control Interval would reduce the I/O processing when using direct access to reference a logical record. The multiple buffers would allow read-ahead (and I/O efficiency) when using sequential processing.

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information to compute the percent of accesses to the data component of the VSAM data set that were direct, using the following algorithm:

$$Percent\ direct\ accesses = \frac{S42AMDRB}{S42AMSRB + S42AMDRB}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces Rule DAS611 when the percent of direct accesses for the data component was greater than the **PCTDIR** guidance variable in USOURCE(DASGUIDE), and the Control Interval size was greater than 4096 bytes. Additionally, CPExpert verifies that the maximum logical records (maximum LRECL) is less than 50% of the Control Interval size. This verification is done to make sure that Rule DAS611 is not produced for VSAM data sets that have spanned records.

The default value for the PCTDIR guidance variable is 80%, so CPExpert will produce Rule DAS611 when more than 80% of the accesses were direct for the data component, and the Control Interval size was more than 4096 bytes.

The following example illustrates the output from Rule DAS611:

```
RULE DAS611:  RELATIVELY LARGE CI SIZE USED FOR DIRECT PROCESSING

   VOLSER: CICS0E.  A relatively large data Control Interval (CI) size
   was used for the VSAM data sets listed below, yet more than 80% of
   the accesses were for direct processing.  You normally should define
   a CI size of 4KB for VSAM data sets used mostly for direct processing.
   The I/O RATE is for the time the data set was open.
                                                                ..      I/O    -ACCESS TYPE (PCT)-    CI
   SMF TIME STAMP   JOB NAME  VSAM DATA SET                    ..       RATE   SEQUENTIAL   DIRECT   SIZE
   17:00,14OCT1997  DFHCOMDS  CICS410.PET.DFHCSD.DATA..................... 174.7       0.0       100.0  18432
```

While not shown in the above example, CPExpert also shows the OPEN time for the VSAM data set. This normally is the duration of the current OPEN. If %LET VSAMSMRY=Y; was specified in USOURCE(DASGUIDE), the OPEN time represents the sum of the times the VSAM data was OPEN for all TYPE64 records in the performance data base.

Suggestion: If CPExpert produces Rule DAS611, you should consider the following alternatives:

- Large Control Interval sizes for direct access normally degrade performance because the entire Control Interval (consisting of multiple logical records) is brought into storage with each access to a record. With direct processing, normally there is a low probability that more than one record in the Control Interval will be referenced. Consequently, you should consider decreasing the Control Interval size.

- If the above action is not appropriate, you can change the **PCTDIR** guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the PCTSEQ guidance variable if you feel that Rule DAS611 is produced prematurely.

- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.5.1.2: Data Control Interval Size

Rule DAS612: Relatively large CI size was used for mixed processing

Finding: CPExpert noticed that a relatively large CI size was used for processing a mixture of sequential and direct accesses. This finding applies only if SMF Type 42 (Data Set Statistics)[1] and SMF Type 64 (VSAM Statistics) records are available in a MXG performance data base.

Impact: This can have a LOW IMPACT, MEDIUM IMPACT, or HIGH IMPACT on the performance of applications referencing the VSAM data. The level of impact depends on the number of I/O operations.

Discussion: A VSAM file structure consists of one or more *Control Intervals (CIs)* and one or more *Control Areas (CAs)*.

- A **Control Interval** is a continuous area of direct access storage that VSAM uses to store logical records. When a logical record is read from direct access storage, the entire Control Interval containing the record is read into a VSAM buffer in virtual storage. The desired logical record is then transferred from the VSAM buffer to a user-defined buffer or work area. While logical records within a Control Interval may vary in length, all Control Intervals in a specific VSAM data set are of the same length.

- A **Control Area** contains one or more Control Intervals. The Control Intervals are grouped together into fixed-length contiguous areas of direct access storage. A VSAM data set is composed of one or more control areas.

Each VSAM data set is defined as a cluster of one or more components.

- The *data component* is the part of a VSAM data set, alternate index, or catalog that contains the data records. The minimum size of a Control Area for a data component is one track, and the maximum size is one cylinder of DASD storage.

- The *index component* is a collection of logically sequenced keys. A key is a value taken from a fixed defined field in each logical record in the VSAM data set. The key identifies the record's position in the data set. Using the index, VSAM is able to randomly retrieve a record from the data component when a request is made for a record with a certain key. The size of the Control Area for an index component is one track of DASD storage.

---

[1] **%LET TYPE42DS = Y;** must be specified in USOURCE(GENGUIDE) must be specified in USOURCE(GENGUIDE) or USOURCE(DASGUIDE) to advise CPExpert that TYPE42DS is available.

The size of Control Intervals can vary from one VSAM data set to another, but all the Control Intervals of a particular data set component must be the same length. Control interval size affects record processing speed and storage requirements:

- Data sets with large control interval sizes require more buffer space in virtual storage.

- Data sets with large control interval sizes require fewer I/O operations to bring a given number of records into virtual storage, because fewer index records must be read.

- Free space is used more efficiently as control interval size increases relative to data record size. This is because there are fewer Control Interval splits and less wasted space.

The type of processing that is used should guide the choice of control interval size, particularly for the data component:

- **Sequential processing**. When sequential processing accounts for most of the accesses, a large Control Interval for the data component would normally be a good choice. This is because multiple records can be read into buffers and processed sequentially. For example, given a 16KB data buffer space, it is better to read two 8 KB Control Intervals with one I/O operation, than to read four 4 KB Control Intervals with two I/O operations.

- **Direct processing**. When direct processing accounts for most of the accesses, a small Control Interval for the data component is preferable. This is because only one logical record is retrieved at a time with direct access. Since a Control Interval normally contains several logical records, a large Control Interval would create unnecessary I/O overhead reading the logical records that would not be accessed. IBM suggests that a 4096 byte Control Interval normally would be appropriate for direct access.

- **Mixed processing**. If the processing is a mixture of sequential and direct, a small Control Interval for the data component with multiple buffers for sequential processing can be a good choice. The small Control Interval would reduce the I/O processing when using direct access to reference a logical record. The multiple buffers would allow read-ahead (and I/O efficiency) when using sequential processing.

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information to compute the percent of sequential accesses to the data component of the VSAM data set, using the following algorithm:

$$\text{Percent sequential accesses} = \frac{S42AMSRB}{S42AMSRB + S42AMDRB}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert also uses the TYPE42DS information to compute the percent of accesses to the VSAM data set that were direct, using the following algorithm:

$$\text{Percent direct accesses} = \frac{S42AMDRB}{S42AMS4B + S42AMDRB}$$

CPExpert produces Rule DAS612 when:

- The percent of direct accesses for the data component was greater than 30%, **and**
- the percent of sequential accesses for the data component was greater than 30%, **and**
- the maximum logical record length was less than 50% of the Control Interval size, **and**
- the Control Interval size was greater than 4096 bytes, **and**
- less than 10 buffers had been allocated.

The result of this algorithm will select those VSAM data sets with data component accessed for both sequential and direct accesses, that can have more than one logical record per Control Interval, and that have not been allocated a relatively large number of buffers.

The following example illustrates the output from Rule DAS612:

```
RULE DAS612:  RELATIVELY LARGE CI SIZE USED FOR MIXED (RANDOM AND SEQ)

   VOLSER: HOLDE2.  A relatively large data Control Interval (CI) size was
   used for the VSAM data sets listed below and few buffers were defined
   for these data sets, yet the data sets were used for a mixture of random
   and sequential processing.  You normally should define a CI size of 4KB
   for these VSAM data sets (for random processing) but define a relatively
   large number of buffers (for sequential processing).   The I/O RATE is
   for the time the data set was open.
                                                                  ..       I/O    -ACCESS TYPE (PCT)-    CI
   SMF TIME STAMP    JOB NAME  VSAM DATA SET                      ..       RATE   SEQUENTIAL   DIRECT   SIZE  BUFFERS
   17:00,14OCT1997   AR170022  IMSDSW4P.DSW.RECON1.DATA....................  32.4      50.0       50.0   32768      2
```

While not shown in the above example, CPExpert also shows the OPEN time for the VSAM data set. This normally is the duration of the current OPEN. If %LET VSAMSMRY=Y; was specified in USOURCE(DASGUIDE), the OPEN time represents the sum of the times the VSAM data was OPEN for all TYPE64 records in the performance data base.

Suggestion: If CPExpert produces Rule DAS612, you should consider the following alternatives:

- Decrease the Control Interval size and increase the number of buffers. As mentioned earlier, IBM suggests that a small data Control Interval with multiple buffers for sequential processing can be a good choice with mixed processing.

- If the above action is not appropriate, you can change the PCTSEQ guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the PCTSEQ guidance variable if you feel that Rule DAS612 is produced prematurely.

- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.5.1.2: Data Control Interval Size

Rule DAS620: The number of data buffers should be increased

Finding: CPExpert noticed that Non-Shared resources (NSR) was specified for VSAM data sets and most of the access was sequential processing. However, relatively few data buffers were assigned to the data sets. Consequently, I/O processing was inefficient. This finding applies only if SMF Type 42 (Data Set Statistics)[1] and SMF Type 64 (VSAM Statistics) records are available in a MXG performance data base.

Impact: This can have a LOW IMPACT, MEDIUM IMPACT, or HIGH IMPACT on the performance of applications referencing the VSAM data. The level of impact depends on the number of direct I/O operations that are used.

Discussion: A VSAM file structure consists of one or more *Control Intervals (CIs)* and one or more *Control Areas (CAs)*.

- A **Control Interval** is a continuous area of direct access storage that VSAM uses to store logical records. When a logical record is read from direct access storage, the entire Control Interval containing the record is read into a VSAM buffer in virtual storage. The desired logical record is then transferred from the VSAM buffer to a user-defined buffer or work area. While logical records within a Control Interval may vary in length, all Control Intervals in a specific VSAM data set are of the same length.

- A **Control Area** contains one or more Control Intervals. The Control Intervals are grouped together into fixed-length contiguous areas of direct access storage. A VSAM data set is composed of one or more Control Areas.

I/O buffers are used by VSAM to read and write control intervals from DASD to virtual storage. For a key-sequenced data set (KSDS) or variable-length relative records data sets (VRRDS), VSAM requires a minimum of three buffers: two buffers for data control intervals[2] and one buffer for an index control interval. Only data buffers are needed for entry-sequenced, for linear data sets (LSDS), and for fixed-length relative record data sets (RRDS).

The VSAM defaults provide these **minimum** buffers. However, to increase performance, there are parameters to override the VSAM default values.

---

[1] **%LET TYPE42DS = Y;** must be specified in USOURCE(GENGUIDE) must be specified in USOURCE(GENGUIDE) or USOURCE(DASGUIDE) to advise CPExpert that TYPE42DS is available.

[2] One of the data buffers is used only for formatting control areas and splitting control intervals and control areas.

Selecting good buffering options can reduce the number of I/O operations, reduce job elapsed time, reduce CPU time, device reduce connect time, and reduce device disconnect time. IBM benchmarks[3] have shown that, depending on the workload, selecting optimal buffering parameters can provide 90% reduction in the number of I/O operations, provide over 65% reduction in job elapsed, provide over 40% reduction in CPU time, and provide over 50% reduction in device connect time. These remarkable savings were achieved simply by altering the buffering characteristics of jobs processing VSAM data sets.

The optimum buffering techniques vary, depending the buffering techniques (or resource pools) used and how the records are accessed (sequential or direct). There are four types of resource pools, depending on the type of data sharing that is implemented:

- **Non-Shared Resource (NSR)**. NSR is the default VSAM buffering technique. With NSR, VSAM buffers are not shared among VSAM data sets, and the buffers are located in the private area. VSAM data sets with NSR buffering can be accessed sequentially or direct (or both). However, NSR is suited for sequential processing because, if the data set access is sequential, the buffers are managed with a read-ahead algorithm. The read-ahead algorithm provides overlap of I/O and CPU processing and is efficient for sequential accesses. Since NSR is oriented toward sequential access, there is no expectation that a record will be re-used (as might exist with direct processing). Consequently, once a record is processed from the NSR buffers, the buffer is likely to be reclaimed for another record read from DASD.

- **Local Shared Resource (LSR)**. With LSR, VSAM buffers normally are shared among VSAM data sets accessed by tasks in the same address space. Since LSR is oriented toward shared (and direct) access, there is an expectation that a record might be re-used. Consequently, buffer management algorithms retain buffers as long as possible, using a least-recently used (LRU) algorithm, after a record is processed from the LSR buffers. There is no read-ahead algorithm with LSR, and there is no inherent overlap of I/O and CPU processing. LSR is appropriate for direct access of VSAM data sets, regardless of whether the data sets are shared.

- **Global Shared Resource (GSR)**. GSR provides serialization of shared resources across multiple systems. In a GSR complex, programs can serialize access to data sets on shared DASD volumes at the *data set level* rather than at the *DASD volume level*. A program on one system can access one data set on a shared volume while other programs on any system can access other data sets on the same volume.

---

[3] *VSAM Demystified*, SG24-6105, Section 2.6.9 (Buffering options)

- **Record Level Sharing (RLS)**. VSAM RLS provides multisystem sharing of VSAM data sets in a parallel sysplex, using cross-system locking. CICS is the exploiter of RLS. RLS enables VSAM data to be shared, with full update capability, between many applications running in many CICS regions. With RLS, CICS regions that share VSAM data sets can reside in one or more MVS images within an MVS parallel sysplex.

As described above, NSR is best used for applications that use sequential or skip sequential as their primary access mode. When data sets are accessed sequentially, performance can be increased by specifying multiple buffers for the data component[4]. When there are multiple data buffers, VSAM uses a read-ahead function to read the next data control intervals into buffers as buffers become available. On the other hand, more buffers than necessary might cause excessive paging or excessive internal processing. There is an optimum point at which more buffers do not continue to improve performance.

Please note that additional buffers do not improve performance if the VSAM data set is processed with direct access. This is because there is no read-ahead I/O activity.

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information to compute the percent of sequential accesses to the VSAM data set, using the following algorithm:

$$\textit{Percent sequential accesses} = \frac{S42AMSRB}{S42AMSRB + S42AMDRB}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces Rule DAS620 when the TYPE42DS S42DSBUF variable showed that NSR was used, the percent of sequential accesses for the data component was greater than 75%, and less than 10 buffers had been assigned to the data component for the VSAM data set.

[4]Having only one I/O buffer for the index component does not hinder performance, because VSAM gets to the next CI by using the horizontal pointers in sequence set records rather than the vertical pointers in the index set. Extra index buffers have little effect during sequential processing.

The following example illustrates the output from Rule DAS620:

```
RULE DAS620:  THE NUMBER OF DATA BUFFERS SHOULD BE INCREASED

   VOLSER: D83NE2.  Non-Shared resources (NSR) was specified for the
   below VSAM data sets and most of the access was sequential processing.
   However, relatively few data buffers were assigned to the data sets.
   You should consider increasing the number of data buffers to at least
   10 buffers, and preferably up to 30 buffers.   The I/O RATE is for the
   time the data set was open.
                                                             ..       I/O    -ACCESS TYPE (PCT)-  BUFFERS
   SMF TIME STAMP    JOB NAME  VSAM DATA SET                 ..       RATE   SEQUENTIAL   DIRECT  ASSIGNED
   17:00,14OCT1997   NETAK31.  OPSYS.NV31.J90.AOC.DSILOGP.DATA............  40.2     100.0        0.0      2
```

While not shown in the above example, CPExpert also shows the OPEN time for the VSAM data set. This normally is the duration of the current OPEN. If %LET VSAMSMRY=Y; was specified in USOURCE(DASGUIDE), the OPEN time represents the sum of the times the VSAM data was OPEN for all TYPE64 records in the performance data base.

Suggestion: If CPExpert produces Rule DAS620, you should consider the following alternatives:

- Increase the number of buffers for the VSAM data sets identified by Rule DAS620. IBM benchmarks[5] have shown that with read sequential access, performance improves significantly if the number of buffers for the data component of VSAM data sets is increased from the default 2 buffers to 10 buffers. Performance continues to increase as more buffers are added, but the improvement was less in IBM's benchmarks, as the number of buffers increased beyond 30 buffers. Consequently, IBM recommends that the number of buffers for the data component of sequentially accessed VSAM data sets be increased from the default to 10 buffers or up to 30 buffers if space permits. This recommendation is only a general guideline, however. The optimum number of data component buffers varies according to the amount of CPU processing[6] done between each read to the data set.

---

[5] *VSAM Demystified*, SG24-6105, Section 2.6.9 (Buffering options)

[6] The amount of CPU processing can be done by the application accessing the VSAM data set, or can be done by work (either system or application work) executing at an equal or higher CPU dispatching priority. The CPU-I/O overlap efficiencies created by multiple VSAM buffers for NSR sequential access occur only when the application otherwise would wait on logical records. Application delay waiting for I/O can be analyzed by the WLM Component of CPExpert, or can be seen in RMF reports for the service class to which the application belongs.

There is one significant exception to this "increase buffers" recommendation; that exception occurs when SHAREOPTIONS 4 has been specified for the VSAM data set.

The SHAREOPTIONS parameter specifies how the component or cluster can be shared among users within one system or across systems. With SHAREOPTIONS 4, the data set can be fully shared by any number of users. This setting does not allow any type of non-RLS access when the data set is already open for RLS processing. With this option, each user is responsible for maintaining both read and write integrity for the data. With SHAREOPTIONS 4, buffers are refreshed at each request. Also, the read-ahead function has no effect and defer write is not used. **Therefore, for SHAREOPTIONS 4, keeping data buffers at a minimum can actually improve performance**.

Unfortunately, SMF information does not describe the SHAREOPTIONS value, so CPExpert is unable to detect that SHAREOPTIONS 4 has been specified for a VSAM data set, and thus cannot suppress this finding for VSAM data sets with SHAREOPTIONS 4.

- Alternatively, you can specify System Managed Buffering (SMB)[7] for the VSAM data sets listed. VSAM can use system-managed buffering to determine the number of buffers and the type of buffer management to use for VSAM data sets. VSAM also determines the number of buffers to locate in Hiperspace for use in direct optimization. To indicate that VSAM is to use SMB, specify either of the following options:

    - Specify the ACCBIAS subparameter of the JCL DD statement AMP parameter and specify **Sequential Optimized (SO)** for the record access bias. This technique provides the most efficient buffers for sequential application. Approximately 500K of processor virtual storage for buffers is required for this technique, defaulted to above 16 MB.

      If 500K processor virtual storage for buffers is not available with the application, consider specifying **Sequential Weighted (SW)** for the record access bias. This technique provides efficient buffers for sequential application. Approximately 100K of processor virtual storage for buffers is required for this technique, defaulted to above 16 MB,.

    - Specify Record Access Bias in the data class and an application processing option in the ACB.

[7] Please review *DFSMS: Using Data Sets* (Section 2.5.4.2.3: Processing Guidelines and Restrictions) before implementing system managed buffering.

For system-managed buffering (SMB), the data set must use both of the following options:

- System Management Subsystem (SMS) storage
- Extended format (DSNTYPE=ext in the data class)

- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.5.5.5.1 Data Buffers for Sequential Access Using Nonshared Resources
Section 2.5.4.2: Tuning for System-Managed Buffering
Section 2.5.4.2.3: Processing Guidelines and Restrictions

IBM Redbook: *VSAM Demystified* (SG24-6105)
Section 2.6.9 (Buffering options)

Finding: CPExpert noticed that direct processing accounted for a significant amount of the I/O activity to the index component of VSAM data sets. However, insufficient buffers were assigned to the index component. The number of index buffers should be increased for optimal performance of the VSAM data sets listed. This finding applies only if SMF Type 42 (Data Set Statistics)[1] and SMF Type 64 (VSAM Statistics) records are available in a MXG performance data base.

Impact: This can have a LOW IMPACT, MEDIUM IMPACT, or HIGH IMPACT on the performance of applications referencing the VSAM data. The level of impact depends on the number of direct I/O operations.

Discussion: A VSAM file structure consists of one or more *Control Intervals (CIs)* and one or more *Control Areas (CAs)*.

- A **Control Interval** is a continuous area of direct access storage that VSAM uses to store logical records. When a logical record is read from direct access storage, the entire Control Interval containing the record is read into a VSAM buffer in virtual storage. The desired logical record is then transferred from the VSAM buffer to a user-defined buffer or work area. While logical records within a Control Interval may vary in length, all Control Intervals in a specific VSAM data set are of the same length.

- A **Control Area** contains one or more Control Intervals. The Control Intervals are grouped together into fixed-length contiguous areas of direct access storage. A VSAM data set is composed of one or more Control Areas.

Each VSAM data set is defined as a cluster of one or more components.

- The *data component* is the part of a VSAM data set, alternate index, or catalog that contains the data records. The minimum size of a Control Area for a data component is one track, and the maximum size is one cylinder of DASD storage.

- The *index component* is a collection of logically sequenced keys. A key is a value taken from a fixed defined field in each logical record in the VSAM data set. The key identifies the record's position in the data set. Using the index, VSAM is able to randomly retrieve a record from the

[1] **%LET TYPE42DS = Y;** must be specified in USOURCE(GENGUIDE) must be specified in USOURCE(GENGUIDE) or USOURCE(DASGUIDE) to advise CPExpert that TYPE42DS is available.

data component when a request is made for a record with a certain key. The size of the Control Area for an index component is one track of DASD storage.

Key-sequenced data sets (KSDS) and variable-length relative records data sets (VRRDS) contain both a data component and an index component. Additionally, each alternate index contains a data component and an index component. Entry-sequenced data sets (ESDS), linear data sets (LDS), and fixed-length relative record data sets (RRDS) contain only a data component.

The index component consists of two parts: *sequence set* and *index set*.

- The sequence set is the lowest level of index control intervals and directly points to the data Control Interval in the data Control Area. There is one Control Interval in the sequence set for each *data* Control Area. This *index* Control Interval contains pointers and high key information for each *data* Control Interval. The index Control Interval also contains horizontal pointers from one sequence set Control Interval to the next higher keyed sequence set Control Interval.

- The records in all levels of the index above the sequence set are called the index set. If there is more than one sequence set Control Interval, VSAM automatically builds another index level. An entry in an index set record consists of the highest possible key in an index record in the next lower level, and a pointer to the beginning of that index record.

I/O buffers are used by VSAM to read and write control intervals from DASD to virtual storage. A minimum of one buffer is required for an index Control Interval. Having only one index I/O buffer does not hinder performance when the VSAM data set is accessed sequentially, because VSAM gets to the next Control Interval by using the horizontal pointers in sequence set records rather than the vertical pointers in the index set.

The minimum of one index buffer[2] is inadequate with **direct** access to the VSAM data set. When using direct access to retrieve a record from a key-sequenced data set or variable-length RRDS (or store a record using keyed access), VSAM needs to examine the index of the data set. When an index record must be retrieved to locate a data record, VSAM makes room for the new index record by deleting the index record that VSAM judges to be least useful under the prevailing circumstances. If only one index buffer were provided, a serious performance problem would occur if an index record were continually deleted from virtual storage to make

[2] Multiple buffers for the *data component* do not increase performance with direct processing, because only one data buffer is used for each access.

room for another index record, and then retrieved again later when it is required.

Providing more than the minimum number of index can significantly improve performance. IBM benchmarks[3] show over 50% reduction in EXCPs and almost 50% reduction in CPU time, by increasing the number of index buffers from 1 to 4 for non-shared resource (NSR) with direct access.

Unused index buffers do not normally degrade performance, so an adequate number should be specified. For optimum performance, the number of index buffers should be at least as large as the number of high-level index set Control Intervals, plus one per string to contain the entire high-level index set and one sequence set control interval per string in virtual storage.

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information (MXG variable DSTYPE) to identify KSDS and VRRDS VSAM data sets. CPExpert selects these VSAM data sets that were used with NSR (using MXG variable S42DSBUF). CPExpert uses the TYPE42DS information to compute the percent of direct accesses to the VSAM data set, using the following algorithm:

$$\textit{Percent direct accesses} = \frac{S42AMDRB}{S42AMSRB + S42AMDRB}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces Rule DAS621 under the following conditions:

- The TYPE42DS S42DSBUF variable showed that NSR was used for KSDS or VRRDS VSAM data sets, **and**

- The percent of direct accesses for the index component was greater than **DIRINDEX** guidance variable in USOURCE(DASGUIDE), **and**

- The number of buffers (the MXG BUFDRNO variable) assigned to the index component was less than the number of index levels (the MXG ACCLEVEL variable) for the VSAM data set.

---

[3] *VSAM Demystified*, SG24-6105, Table 6 (NSR buffering with direct access - STRNO=1)

The default value for the DIRINDEX guidance variable is 25%, so CPExpert will produce Rule DAS621 for NSR VSAM data sets when more than 25% of the accesses were direct for the index component and the number of buffers assigned to the index component was less than the number of index levels.

The following example illustrates the output from Rule DAS621:

```
RULE DAS621:  THE NUMBER OF INDEX BUFFERS SHOULD BE INCREASED

   VOLSER: PRD005.  Non-Shared resources (NSR) was specified as the
   buffering technique for the below VSAM data sets, and most of the
   accesses were random processing.  However, relatively few index
   buffers were assigned to the data sets.  You should consider
   increasing the number of index buffers to the number of index
   levels.  The I/O RATE is for the time the data set was open.
                                                                ..       I/O    -ACCESS TYPE (PCT)-  INDEX    BUFFERS
   SMF TIME STAMP    JOB NAME  VSAM DATA SET                    ..       RATE   SEQUENTIAL   DIRECT  LEVELS   ASSIGNED
   10:04,19SEP2002   PGRED01D  PWFWFA.PK.GENTRAN.EDIOEA.INDEX............. 30.7       0.0      100.0      3         1
   10:20,19SEP2002   PFFED01D  PWFWFA.PK.GENTRAN.EDIOEA.INDEX............. 15.7       0.0      100.0      3         1
```

While not shown in the above example, CPExpert also shows the OPEN time for the VSAM data set. This normally is the duration of the current OPEN. If %LET VSAMSMRY=Y; was specified in USOURCE(DASGUIDE), the OPEN time represents the sum of the times the VSAM data was OPEN for all TYPE64 records in the performance data base.

Suggestion: If CPExpert produces Rule DAS621, you should consider the following alternatives:

- **Increase the number of index buffers**. You should consider increasing the number of index buffers allocated for the VSAM data sets listed by Rule DAS621.

- **Consider the use of System Managed Buffering**. If the data sets listed by Rule DAS621 are SMS managed, have extended format, and your installation has DFSMS V1R4 or later, you can use system managed buffering (SMB)[4]. System managed buffering enables VSAM to determine the optimum number of index and data buffers, as well as the type of buffer management (LSR or NSR). IBM benchmarks[5] show up to

[4] Please review *DFSMS: Using Data Sets* (Section 2.5.4.2.3: Processing Guidelines and Restrictions) before implementing system managed buffering.

[5] *VSAM Demystified*, SG24-6105, Table 7 (Direct access: Benefits of using SMF: Updates and insertions)

90% reduction in EXCPs, DASD connect time, and CPU time when converting to system managed buffering!

To indicate that VSAM is to use SMB, specify either of the following options:

- Specify the ACCBIAS subparameter of the JCL DD statement AMP parameter.

- Specify Record Access Bias in the data class and an application processing option in the ACB. The ACB must be NSR and MACRF cannot contain any of the following:

    - ICI (Improved control interval processing)
    - AIX Processing the data set through the alternate index of the path specified in the DDname
    - UBF (Management of I/O buffers left up to the VSAM user)

Regardless of whether the JCL DD statement AMP is used or the data class/ACB is used, you should specify the following for the record access bias:

- **Specify Direct Optimized (DO)** for the record access bias for the record access bias if *direct access is near 100% for the VSAM data sets listed*. The DO processing technique optimizes for totally random record access. This technique overrides the user specification for non-shared resources (NSR) buffering with a local shared resources (LSR) implementation of buffering.

- **Specify Direct Weighted (DW)** for the record access bias if the data sets are processed using a mixture of direct and sequential access, and *direct access is dominate*. DW processing provides the minimum read-ahead buffers for sequential retrieval and the maximum index buffers for direct requests.

- **Specify Sequential Weighted (SW)** for the record access bias if the data sets are processed using a mixture of direct and sequential access, and *sequential access is dominate*. This technique uses read-ahead buffers for sequential requests and provides additional index buffers for direct requests. Approximately 100K of processor virtual storage for buffers is required for this technique, defaulted to above 16 MB.

- **Consider the use of Batch LSR**. For batch applications **and if Rule DAS621 shows that direct processing[6] is near 100%**, you can use Batch LSR. Batch LSR provides advantages in an application using VSAM NSR buffering techniques to switch to LSR without changing the application source code or link-editing the application again. Only a JCL change is required.

- If the above actions are not appropriate, you can change the **DIRINDEX** guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the DIRINDEX guidance variable if you feel that Rule DAS621 is produced prematurely.

- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.5.5.5.1 Data Buffers for Sequential Access Using Nonshared Resources
Section 2.5.4.2: Tuning for System-Managed Buffering

IBM Redbook: *VSAM Demystified* (SG24-6105)
Section 2.6.9 (Buffering options)

[6] Using the Batch LSR subsystem with sequential access could degrade performance rather than improve it. This is because Batch LSR does not provide "read-ahead" capability. The "read-ahead" capability is essential for good performance with sequential access.

Rule DAS622: The number of index buffers should be increased (STRNO value)

Finding: The number of index buffers was not greater than the number of strings specified in the Access Control Block (ACB) STRNO value. The number of index buffers should be increased for optimal performance of the VSAM data sets listed. This finding applies only if SMF Type 42 (Data Set Statistics)[1] and SMF Type 64 (VSAM Statistics) records are available in a MXG performance data base.

Impact: This can have a LOW IMPACT, MEDIUM IMPACT, or HIGH IMPACT on the performance of applications referencing the VSAM data. The level of impact depends on the number of direct I/O operations.

Discussion: A VSAM file structure consists of one or more *Control Intervals (CIs)* and one or more *Control Areas (CAs)*.

- A **Control Interval** is a continuous area of direct access storage that VSAM uses to store logical records. When a logical record is read from direct access storage, the entire Control Interval containing the record is read into a VSAM buffer in virtual storage. The desired logical record is then transferred from the VSAM buffer to a user-defined buffer or work area. While logical records within a Control Interval may vary in length, all Control Intervals in a specific VSAM data set are of the same length.

- A **Control Area** contains one or more Control Intervals. The Control Intervals are grouped together into fixed-length contiguous areas of direct access storage. A VSAM data set is composed of one or more Control Areas.

Each VSAM data set is defined as a cluster of one or more components.

- The *data component* is the part of a VSAM data set, alternate index, or catalog that contains the data records. The minimum size of a Control Area for a data component is one track, and the maximum size is one cylinder of DASD storage.

- The *index component* is a collection of logically sequenced keys. A key is a value taken from a fixed defined field in each logical record in the VSAM data set. The key identifies the record's position in the data set. Using the index, VSAM is able to randomly retrieve a record from the data component when a request is made for a record with a certain key.

[1] **%LET TYPE42DS = Y;** must be specified in USOURCE(GENGUIDE) must be specified in USOURCE(GENGUIDE) or USOURCE(DASGUIDE) to advise CPExpert that TYPE42DS is available.

The size of the Control Area for an index component is one track of DASD storage.

Key-sequenced data sets (KSDS) and variable-length relative records data sets (VRRDS) contain both a data component and an index component. Additionally, each alternate index contains a data component and an index component. Entry-sequenced data sets (ESDS), linear data sets (LDS), and fixed-length relative record data sets (RRDS) contain only a data component.

The index component consists of two parts: *sequence set* and *index set*.

- The sequence set is the lowest level of index control intervals and directly points to the data Control Interval in the data Control Area. There is one Control Interval in the sequence set for each *data* Control Area. This *index* Control Interval contains pointers and high key information for each *data* Control Interval. This index Control Interval also contains horizontal pointers from one sequence set Control Interval to the next higher keyed sequence set Control Interval.

- The records in all levels of the index above the sequence set are called the index set. If there is more than one sequence set Control Interval, VSAM automatically builds another index level. An entry in an index set record consists of the highest possible key in an index record in the next lower level, and a pointer to the beginning of that index record.

I/O buffers are used by VSAM to read and write control intervals from DASD to virtual storage. A minimum of one buffer is required for an index Control Interval. Having only one index I/O buffer does not hinder performance when the VSAM data set is accessed sequentially, because VSAM gets to the next Control Interval by using the horizontal pointers in sequence set records rather than the vertical pointers in the index set.

A string is a request to a VSAM data set requiring data set positioning. If different *concurrent* accesses to the same data set are necessary, multiple strings are used. If multiple strings are used, each string requires exclusive control of an index I/O buffer. Therefore, the value specified for the STRNO parameter (in the ACB or GENCB macro, or AMP parameter) is the minimum number of index I/O buffers required when requests that require concurrent positioning are issued.

If the number of I/O buffers provided for index records is greater than the number of requests that require concurrent positioning (specified by the STRNO parameter), one buffer is used for the highest-level index record. Any additional buffers are used, as required, for other index-set index records. With direct access, the minimum number of index buffers should be one more than the value of the STRNO if VSAM is to keep the highest-level index record resident. Keeping the highest-level index record resident can significantly improve performance, at a modest increase in virtual storage used for index buffers

For optimum performance, the number of index buffers should be at least as large as the number of high-level index set Control Intervals, **plus one per string to contain the entire high-level index set and one sequence set control interval per string in virtual storage.**

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information to compute the percent of accesses to the VSAM data set that were CPExpert uses the TYPE42DS information to compute the percent of direct accesses to the VSAM data set, using the following algorithm:

$$Percent\ direct\ accesses = \frac{S42AMDRB}{S42AMSRB + S42AMDRB}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces Rule DAS622 under the following conditions:

- The TYPE42DS S42DSBUF variable showed that NSR was used for KSDS or VRRDS VSAM data sets, **and**

- The percent of direct accesses for the index component was greater than **DIRINDEX** guidance variable in USOURCE(DASGUIDE), **and**

- The STRNO specification in the ACB (the MXG ACBSTRNO) was greater than one (indicating that concurrent accesses had been specified for direct processing), **and**

- The number of buffers (the MXG BUFDRNO variable) assigned to the index component was less than the ACBSTRNO value, plus 1.

The default value for the DIRINDEX guidance variable is 25%, so CPExpert will produce Rule DAS622 for NSR VSAM data sets when more than 25% of the accesses were direct for the index component and the number of buffers assigned to the index component was less than the STRNO value, plus 1.

The following example illustrates the output from Rule DAS622:

```
RULE DAS622:  THE NUMBER OF INDEX BUFFERS SHOULD BE INCREASED (STRNO VALUE)

   VOLSER: MVS902.  Non-Shared resources (NSR) was specified as the
   buffering technique for the below VSAM data sets, and most of the
   access were direct processing. However, relatively few index buffers
   were assigned to the data sets. You should consider increasing the
   number of index buffers to 1 more than the number of strings specified
   in the STRNO in the ACB. The I/O RATE is for the time the data set was
   open.
                                                                ..       I/O    -ACCESS TYPE (PCT)-          BUFFERS
   SMF TIME STAMP    JOB NAME  VSAM DATA SET                    ..       RATE   SEQUENTIAL   DIRECT  STRINGS ASSIGNED
   10:10,19SEP2002   NDMMON..  SDPDPA.PK.MVSP.RT.NDMTCF.INDEX.............  39.2       0.0      100.0      5        5
   10:10,19SEP2002   PCGORDER  SDPDPA.PK.MVSP.RT.NDMDRD.INDEX.............  12.1       0.0      100.0      3        3
   10:10,19SEP2002   NDMMON..  SDPDPA.PK.MVSP.RT.NDMTCF.INDEX............. 242.8       0.0      100.0      5        5
   10:14,19SEP2002   PGDOS06R  SDPDPA.PK.MVSP.CS.NDMDRD.INDEX.............  12.9       0.0      100.0      3        3
   10:14,19SEP2002   PGDOS06R  SDPDPA.PK.MVSP.CS.NDMDRD.INDEX.............  16.5       0.0      100.0      3        3
```

While not shown in the above example, CPExpert also shows the OPEN time for the VSAM data set. This normally is the duration of the current OPEN. If %LET VSAMSMRY=Y; was specified in USOURCE(DASGUIDE), the OPEN time represents the sum of the times the VSAM data was OPEN for all TYPE64 records in the performance data base.

Suggestion: If CPExpert produces Rule DAS622, you should consider the following alternatives:

- **Increase the number of buffers**. You should consider increasing the number of index buffers for the VSAM data sets identified by Rule DAS622 to be one greater than the number of concurrent accesses (as specified by the STRNO parameter).

- Alternatively, you can specify System Managed Buffering (SMB)[2] for the VSAM data set. VSAM can use system-managed buffering to determine the number of buffers and the type of buffer management to use for VSAM data sets. VSAM also determines the number of buffers to locate in Hiperspace for use in direct optimization. To indicate that VSAM is to use SMB, specify either of the following options:

  - Specify the ACCBIAS subparameter of the JCL DD statement AMP parameter and specify Sequential Optimized (SO) for the record access bias.

  - Specify Record Access Bias in the data class and an application processing option in the ACB.

[2] Please review *DFSMS: Using Data Sets* (Section 2.5.4.2.3: Processing Guidelines and Restrictions) before implementing system managed buffering.

For system-managed buffering (SMB), the data set must use both of the following options:

- System Management Subsystem (SMS) storage
- Extended format (DSNTYPE=ext in the data class)

- If the above actions are not appropriate, you can change the **DIRINDEX** guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the DIRINDEX guidance variable if you feel that Rule DAS622 is produced prematurely.

- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.5.4.4.2: Index Buffers for Direct Access
Section 2.5.4.2: Tuning for System-Managed Buffering

IBM Redbook: *VSAM Demystified* (SG24-6105)
Section 2.6.9 (Buffering options)

Rule DAS625: NSR was used, but a large percent of the access was direct

Finding: CPExpert noticed that Non-shared resources (NSR) was used as the access method for the VSAM data sets listed, but most of the accesses were direct. This finding applies only if SMF Type 42 (Data Set Statistics) information is available[1] in a MXG performance data base

Impact: This can have a LOW IMPACT, MEDIUM IMPACT, or HIGH IMPACT on the performance of applications referencing the VSAM data. The level of impact depends on the number of direct I/O operations.

Discussion: I/O buffers are used by VSAM to read and write control intervals from DASD to virtual storage. Two buffering techniques are used with VSAM data sets that are accessed only on a locaL[2] system: Non-shared resource (NSR) and Local shared resource (LSR).

- **Non-Shared Resource (NSR)**. NSR is the default VSAM buffering technique. With NSR, VSAM buffers are not shared among VSAM data sets, and the buffers are located in the private area. VSAM data sets with NSR buffering can be accessed sequentially or direct (or both). However, NSR is suited for sequential processing because, if the data set access is sequential, the buffers are managed with a read-ahead algorithm. The read-ahead algorithm provides overlap of I/O and CPU processing and is efficient for sequential accesses. Since NSR is oriented toward sequential access, there is no expectation that a record will be re-used (as might exist with direct processing). Consequently, once a record is processed from the NSR buffers, the buffer is likely to be reclaimed for another record read from DASD.

- **Local Shared Resource (LSR)**. With LSR, VSAM buffers normally are shared among VSAM data sets accessed by tasks in the same address space. Since LSR is oriented toward shared (and direct) access, there is an expectation that a record might be re-used. Consequently, buffer management algorithms retain buffers as long as possible, using a least-recently used (LRU) algorithm, after a record is processed from the LSR buffers. There is no read-ahead algorithm with LSR, and there is no inherent overlap of I/O and CPU processing. LSR is appropriate for

[1] **%LET TYPE42DS = Y;** must be specified in USOURCE(GENGUIDE) must be specified in USOURCE(GENGUIDE) or USOURCE(DASGUIDE) to advise CPExpert that TYPE42DS is available.

[2] Two other techniques can be used: global shared resource (GSR) and record level sharing (RLS). GSR provides serialization of shared resources across multiple systems. VSAM RLS provides multisystem sharing of VSAM data sets in a parallel sysplex.

direct access of VSAM data sets, regardless of whether the data sets are shared.

As described above, NSR is best used for applications that use sequential or skip sequential as their primary access mode. NSR is not suited for direct processing, although NSR often **is** used for direct processing because it is easy to use and is the default buffering technique. Nonetheless, performance can be significantly improved if LSR is used for direct processing of VSAM data sets.

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information to compute the percent of direct accesses to the VSAM data set, using the following algorithm:

$$\textit{Percent direct accesses} = \frac{\textit{S42AMDRB}}{\textit{S42AMSRB} + \textit{S42AMDRB}}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces Rule DAS625 under the following conditions:

- The TYPE42DS S42DSBUF variable showed that NSR was used for KSDS or VRRDS VSAM data sets, and

- The percent of direct accesses for the data component was greater than the value specified for the NSRDIR guidance variable in USOURCE(DASGUIDE.

The default value for the NSRDIR guidance variable is 75%, so CPExpert will produce Rule DAS625 when NSR was specified as the buffering technique, and more than 75% of the accesses were direct for the data component.

The following example illustrates the output from Rule DAS625:

```
RULE DAS625:  NSR WAS USED, BUT LARGE PERCENT OF ACCESS WAS DIRECT

   VOLSER: MVS902.  Non-Shared resources (NSR) was specified as the
   buffering technique for the below VSAM data sets, but more than 75%
   of the I/O activity was direct access.  NSR is not designed for direct
   access, and many of the advantages of NSR are not available for direct
   access. You should consider Local Shared Resources (LSR) for the below
   VSAM data sets (perhaps using System Managed Buffers to facilitate the
   use of LSR).  The I/O RATE is for the time the data set was open. The
   SMF TIME STAMP and JOB NAME are from the last record for the data set.
                                                                    ..        I/O     OPEN    -ACCESS TYPE (PCT)-
   SMF TIME STAMP    JOB NAME  VSAM DATA SET                        ..        RATE  DURATION  SEQUENTIAL   DIRECT
   13:19,19SEP2002   NRXX807.  SDPDPA.PK.MVSP.RT.NDMGIX.DATA...............  8.4   0:07:08       0.0      100.0
   13:19,19SEP2002   NRXX807.  SDPDPA.PR.MVSP.RT.NDMGIXD.DATA.............  11.2   0:06:42       0.0      100.0
   13:33,19SEP2002   TSJHM...  SDPDPA.PR.MVSP.RT.NDMRQFDA.DATA.............  0.3   2:21:58       0.0      100.0
   13:33,19SEP2002   TSJHM...  SDPDPA.PR.MVSP.RT.NDMRQF.DATA...............  2.8   3:37:53       0.0      100.0
   13:33,19SEP2002   TSJHM...  SDPDPA.PK.MVSP.RT.NDMTCF.DATA...............  11.1   6:24:10       0.1       99.9
```

While not shown in the above example, CPExpert also shows the OPEN time for the VSAM data set. This normally is the duration of the current OPEN. If %LET VSAMSMRY=Y; was specified in USOURCE(DASGUIDE), the OPEN time represents the sum of the times the VSAM data was OPEN for all TYPE64 records in the performance data base.

Suggestion: If CPExpert consistently produces Rule DAS625 for the same VSAM data sets, and the related jobs are having performance problems randomly accessing VSAM data, *you can improve performance with no changes in your applications*. You should consider the following alternatives:

- **Consider the use of System Managed Buffering**. If the data sets listed by Rule DAS625 are SMS managed, have extended format, and your installation has DFSMS V1R4 or later, you can use system managed buffering[3]. System managed buffering enables VSAM to determine the optimum number of index and data buffers, as well as the type of buffer management (LSR or NSR). IBM benchmarks[4] show up to 90% reduction in EXCPs, DASD connect time, and CPU time when converting to system managed buffering!

  To indicate that VSAM is to use SMB, specify either of the following options:

  - Specify the ACCBIAS subparameter of the JCL DD statement AMP parameter.

[3] Please review *DFSMS: Using Data Sets* (Section 2.5.4.2.3: Processing Guidelines and Restrictions) before implementing system managed buffering.

[4] *VSAM Demystified*, SG24-6105, Table 7 (Direct access: Benefits of using SMF: Updates and insertions)

- Specify Record Access Bias in the data class and an application processing option in the ACB. The ACB must be NSR and MACRF cannot contain any of the following:

  - ICI (Improved control interval processing)
  - AIX Processing the data set through the alternate index of the path specified in the DDname
  - UBF (Management of I/O buffers left up to the VSAM user)

- **Specify Direct Optimized (DO)**. Regardless of whether the JCL DD statement AMP is used or the data class/ACB is used, you should specify Direct Optimized (DO) for the record access bias if *direct access is near 100% for the VSAM data sets listed*. The DO processing technique optimizes for totally random record access. This technique overrides the user specification for nonshared resources (NSR) buffering with a local shared resources (LSR) implementation of buffering.

- **Specify Direct Weighted (DW)** for the record access bias if the data sets are processed using a mixture of direct and sequential access, but direct access is not near[5] 100%. DW processing provides the minimum read-ahead buffers for sequential retrieval and the maximum index buffers for direct requests.

- **Consider the use of Batch LSR**. For batch applications **and if Rule DAS625 shows that direct processing[6] is near 100%**, you can use Batch LSR. Batch LSR provides advantages in an application using VSAM NSR buffering techniques to switch to LSR without changing the application source code or link-editing the application again. Only a JCL change is required.

- If the above actions are not appropriate, you can change the NSRDIR guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the NSRDIR guidance variable if you feel that Rule DAS625 is produced prematurely.

- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential

---

[5]Note that the default value for the NSRDIR guidance variable is 75, indicating that more than 75% of the processing for the VSAM data set was direct access. Unless this guidance is reduced, Rule DAS625 should not be produced unless more than 75% of the accesses were direct.

[6]Using the Batch LSR subsystem with sequential access could degrade performance rather than improve it. This is because Batch LSR does not provide "read-ahead" capability. The "read-ahead" capability is essential for good performance with sequential access.

VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.5.4: Determining I/O Buffer Space for Nonshared Resource
Section 2.5.4.2.3: Processing Guidelines and Restrictions

Rule DAS635: LSR was used, but a large percent of the access was sequential

Finding: CPExpert noticed that local shared resources (LSR) was used as the access method for the VSAM data sets listed, but most of the accesses were sequential. This finding applies only if SMF Type 42 (Data Set Statistics) information is available[1] in a MXG performance data base

Impact: This can have a LOW IMPACT, MEDIUM IMPACT, or HIGH IMPACT on the performance of applications referencing the VSAM data. The level of impact depends on the number of sequential I/O operations.

Discussion: I/O buffers are used by VSAM to read and write control intervals from DASD to virtual storage. Two buffering techniques are used with VSAM data sets that are accessed only on a locaL[2] system: Non-shared resource (NSR) and Local shared resource (LSR).

- **Non-Shared Resource (NSR)**. NSR is the default VSAM buffering technique. With NSR, VSAM buffers are not shared among VSAM data sets, and the buffers are located in the private area. VSAM data sets with NSR buffering can be accessed sequentially or direct (or both). However, NSR is suited for sequential processing because, if the data set access is sequential, the buffers are managed with a read-ahead algorithm. The read-ahead algorithm provides overlap of I/O and CPU processing and is efficient for sequential accesses. Since NSR is oriented toward sequential access, there is no expectation that a record will be re-used (as might exist with direct processing). Consequently, once a record is processed from the NSR buffers, the buffer is likely to be reclaimed for another record read from DASD.

- **Local Shared Resource (LSR)**. With LSR, VSAM buffers normally are shared among VSAM data sets accessed by tasks in the same address space. Since LSR is oriented toward shared (and direct) access, there is an expectation that a record might be re-used. Consequently, buffer management algorithms retain buffers as long as possible, using a least-recently used (LRU) algorithm, after a record is processed from the LSR buffers. There is no read-ahead algorithm with LSR, and there is no inherent overlap of I/O and CPU processing. LSR is appropriate for

[1] **%LET TYPE42DS = Y;** must be specified in USOURCE(GENGUIDE) must be specified in USOURCE(GENGUIDE) or USOURCE(DASGUIDE) to advise CPExpert that TYPE42DS is available.

[2] Two other techniques can be used: global shared resource (GSR) and record level sharing (RLS). GSR provides serialization of shared resources across multiple systems. VSAM RLS provides multisystem sharing of VSAM data sets in a parallel sysplex.

direct access of VSAM data sets, regardless of whether the data sets are shared.

Selecting good buffering options can reduce the number of I/O operations, reduce job elapsed time, reduce CPU time, device reduce connect time, and reduce device disconnect time. IBM benchmarks[3] have shown that, depending on the workload, selecting optimal buffering parameters can provide 90% reduction in the number of I/O operations, provide over 65% reduction in job elapsed, provide over 40% reduction in CPU time, and provide over 50% reduction in device connect time. These remarkable savings were achieved simply by altering the buffering characteristics of jobs processing VSAM data sets.

As described above, **NSR** is best used for applications that use sequential or skip sequential as their primary access mode. When data sets are accessed sequentially, performance can be increased by specifying multiple data buffers. When there are multiple data buffers, VSAM uses a read-ahead function to read the next data control intervals into buffers as buffers become available. On the other hand, more buffers than necessary might cause excessive paging or excessive internal processing. There is an optimum point at which more buffers do not continue to improve performance.

LSR is not suited for applications that use sequential or skip sequential as their primary access mode, because there is no read-ahead algorithm with LSR, and there is no inherent overlap of I/O and CPU processing. Consequently, using LSR for sequential access processing could degrade rather than improve performance.

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPExpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPExpert uses the TYPE42DS information to compute the percent of sequential accesses to the VSAM data set, using the following algorithm:

$$\textit{Percent sequential accesses} = \frac{S42AMSRB}{S42AMSRB + S42AMDRB}$$

where: S42AMSRB = Blocks read using sequential access
S42AMDRB = Blocks read using direct access

CPExpert produces Rule DAS635 under the following conditions:

---

[3] *VSAM Demystified*, SG24-6105, Section 2.6.9 (Buffering options)

- The TYPE42DS S42DSBUF variable showed that LSR was used for KSDS or VRRDS VSAM data sets, and

- The percent of sequential accesses for the data component was greater than the value specified for the LSRSEQ guidance variable in USOURCE(DASGUIDE.

The default value for the LSRSEQ guidance variable is 75%, so CPExpert will produce Rule DAS635 when LSR was specified as the buffering technique, and more than 75% of the accesses were sequential for the data component.

The following example illustrates the output from Rule DAS635:

```
RULE DAS635:  LSR WAS USED, BUT LARGE PERCENT OF ACCESS WAS SEQUENTIAL

   VOLSER: PRD001.  Local shared resources (LSR) was specified as the
   buffering technique for the VSAM data sets listed below, but more than
   75% of the I/O activity was sequential access.  LSR is not designed for
   sequential access, and many of the advantages of LSR are not available
   for sequential access.  You should consider Non-Shared Resources (NSR)
   for the below VSAM data sets.  The I/O RATE is for the time the data set
   was open. The SMF TIME STAMP and JOB NAME are from the last record for
   the data set.
                                                             ..       I/O     OPEN    -ACCESS TYPE (PCT)-
   SMF TIME STAMP    JOB NAME  VSAM DATA SET                 ..       RATE  DURATION  SEQUENTIAL   DIRECT
   10:36,19SEP2002   PICBUYD1  PNPFFA.PK.INFOREM.IC.ROD.DATA...............   1.6   0:02:13    100.0        0.0
```

While not shown in the above example, CPExpert also shows the OPEN time for the VSAM data set. This normally is the duration of the current OPEN. If %LET VSAMSMRY=Y; was specified in USOURCE(DASGUIDE), the OPEN time represents the sum of the times the VSAM data was OPEN for all TYPE64 records in the performance data base.

Suggestion: If CPExpert consistently produces Rule DAS635 for the same VSAM data sets, and the related jobs are having performance problems accessing VSAM data, you should consider the following alternatives:

- Convert the VSAM buffer technique from LSR to Non-shared resource (NSR) for the VSAM data sets. This will allow VSAM to manage the buffers with a read-ahead algorithm.   The read-ahead algorithm will provide overlap of I/O and CPU processing and is efficient for sequential accesses.

- If the above action is not appropriate, you can change the LSRSEQ guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the LSRSEQ guidance variable if you feel that Rule DAS635 is produced prematurely.

- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: IBM Redbook: *VSAM Demystified* (SG24-6105)
Section 2.6.9 (Buffering options)

Your turn:

This manual has described how to use the DASD Component to analyze performance constraints with your overall computing environment.

We would appreciate receiving any comments you have regarding this document (style, content, clarity, etc.), or suggestions for improving the DASD Component (ease of use, new rules, changes to rules, etc.). Please send your comments to:

Don Deese
Computer Management Sciences, Inc.
6076-D Franconia Road
Alexandria, VA 22310
www.cpexpert.com

Comments: